

STIC-Biotech/ChemLib

112172

From: Ramirez, Delia  
Sent: Wednesday, January 14, 2004 5:38 PM  
To: STIC-Biotech/ChemLib  
Subject: case 09/847081

Hi,

I would like to request the following interference searches: seq id 1 and 2 in the nucleic acid databases.

Thank you,

Delia M. Ramirez, Ph.D.  
Patent Examiner  
Recombinant Enzymes-Art Unit 1652  
USPTO  
1911 S. Clark Street, Crystal Mall # 10D06, Mail room 10D01  
Arlington, VA 22202  
(703) 306-0288  
delia.ramirez@uspto.gov

RECEIVED  
JAN 15 2004  
(STIC)

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 1/15/04  
Date Completed: 1/16/04  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH: /  
NA Sequences: \_\_\_\_\_  
AA Sequences: / - review to NA  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \$34  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_









```
/cissue_type="leaf"
/dev_stage="Mature leaf"
1..1119
/gene="PSY2"
1..934
/partial
/gene="PSY2"
/notes="RT-PCR product"
/codon_start=2
/product="phytoene synthase"
/protein_id="AAA34187.1"
/db_xref="GI:437020"
/translation="DPDIVLPGNLGLSEAYDRCEVCAEVAKTFYLTGMTLMTPPRR
AIAIYVWCRRDELVDGPNASHITPQALDREARLEDEIENRPFMDLAALSDTVSR
PEVDIOPEDWDEGMKSRNNFDELILYCYVAGTGLMSPVPIGIAPEKAT
TESVYNAALAIAGIANTLILADVGEDADRGHVLPODLAAGLSDIEDIENRKYTDK
WRIPMKOIQRARKEFDEAEKVTELSASRNPVLASLLIKIIDEIENDYNNFTR
RAYVSKPKKLLTLPAYARSLVPPKSTISCPAKT"
translc_peptide 5..76
/partial
/gene="PSY2"
/notes="approximate location"
mat_peptide
77..931
/gene="PSY2"
/product="phytoene synthase"
BASE COUNT 343 a 198 c 269 g 309 t
ORIGIN
Query Match 49.9%; Score 862.8; DB 8; Length 1119;
Best Local Similarity 89.0%; Pred. No. 5.9e-200;
Matches 984; Conservative 0; Mismatches 102; Indels 20; Gaps 4;
QY 633 GAAGCGGATATGTTGTTCCAGGGAAATTTGGGCTTTGGAGTGAAGCATATGATGCTTG 692
Db 1 GGAATCGGATATGTTGTTCCGCGTATTTGGGCTTTGGAGTGAAGCATATGATGCTTG 60
QY 693 TGGCGAAGTATGTCAGAGTATGCAAGAATTTTACTTGAAGAACCACTAATGACCCC 752
Db 61 TGGCGAAGTATGTCAGAGTATGCAAGAATTTTACTTGAAGAACCACTAATGATGCTTC 120
QY 753 AGAGAGAGAGAGCTATCTGGGCAATATATGTGTGTGTCAGAGAGAACGATGAGCTTGT 812
Db 121 AGACAGAGAGAGAGCTATCTGGGCAATATATGTGTGTGTCAGAGAGAACGATGAGCTTGT 180
QY 813 TGATGGCCCTAATGATCCACATTAATCTCGCAAGCTTTAGATGATGAGTGAAGACACAGGCT 872
Db 181 TGATGGCCCTAATGATCCACATTAATCTCGCAAGCTTTAGATGATGAGTGAAGACACAGGCT 240
QY 873 GGAAGATATTTTCAGTGGCGGCGCATTTGATATGCTTGATGCTGCTTTATCCGATCTGT 932
Db 241 GGAAGATATTTTCAGTGGCGGCGCATTTGATATGCTTGATGCTGCTTTATCCGACACTGT 300
QY 933 CTCGAGATTTCTGTTGATATTCAGCATTCAGAGATATGATTTGAAGAAATCGTATGGA 992
Db 301 TTCGAGATTTCTGTTGATATTCAGCATTCAGAGATATGATTTGAAGAAATCGTATGGA 360
QY 993 CTTTGGAAATCCAGATATAAACTTTCAGATGAGCATATCTCTATGTTTACTATGCTTGC 1052
Db 361 CTTTGGAAATCCAGATATAAACTTTCAGATGAGCATATCTCTATGTTTACTATGCTTGC 420
QY 1053 TGGTACTGATGATGATGATGCTTCAGATTTAGGATTTAGCACTGATCAAAAGCAAC 1112
Db 421 TGGTACTGATGATGATGATGCTTCAGATTTAGGATTTAGCACTGATCAAAAGCAAC 480
QY 1113 AACAGAGATGATATATATGCTTGGCTTTAGGCTTGAATCAATCAACCAATAT 1172
Db 481 AACAGAGATGATATATATGCTTGGCTTTAGGCTTGAATCAATCAACCAATAT 540
QY 1173 ACTCAGAGATGATGATGATGCTTGGCTTGAAGAGAGAGAGATATCTTGGCTCAAGATCAAT 1232
Db 541 ACTCAGAGATGATGATGATGCTTGGCTTGAAGAGAGAGATATCTTGGCTCAAGATCAAT 600
QY 1233 AGCAGAGGAGGCTCTCCGAGAGACATATTTGCTGGAAGATGATGATGAGAG 1292
```

```
|||||
Db 601 AGCAGAGGAGGCTCTCCGAGAGACATATTTTCTGGAAAGGAGTCAATGATGAG 660
QY 1293 GAACCTTTGAAAGAAACAATTCAGAGGCGCAGAGAAATTTCTTGTATGATCAGAGAAAG 1352
Db 661 AATCTTTATGAAAGAAATTCAGAGGCGCAGAGAAATTTCTTGTATGATGAGAGAAAG 720
QY 1353 TGTCAAGAACTGCACTGCTAGTATGATGATGCTGTGTTAAGAGGCTGCTGTGTTATCG 1412
Db 721 TGTCAAGAACTGCACTGCTAGTATGATGATGCTGTGTTAAGAGGCTGCTGTGTTATCG 780
QY 1413 CAAGATATTTGACGAGATTTGAAGCCAGACTATCAACAACCTTCACAGAGAGGCTTATGT 1472
Db 781 CAAGATATTTGACGAGATTTGAAGCCAGACTATCAACAACCTTCACAGAGAGGCTTATGT 840
QY 1473 TAGCAGCCAAAGAAAGCTTTCACCTTTCAGCTTTCCTTATGCAAAATCTCTTTCGCCCC 1532
Db 841 GAGCAAGCCAAAGAAAGCTTTCAGCTTTCAGCTTTCCTTATGCAAGATCTTAGTGCCCC 900
QY 1533 TAATAGAACTTCTCTCCACTAGCAAGAAGCATGATGATGATGAGT-----CAA 1584
Db 901 TAGTCAAACTTCTCTCCACTAGCAAGAAGCATGATGATGATGATGAGT-----CAA 1641
QY 1585 TGAGTATTTACACTTAAGAA---ACTCAGGATCTTGAATGATGATGATGATGATGATGATGAT 1641
Db 961 GCTATACATACACTTAAGAAAGAAAGCATGATGATGATGATGATGATGATGATGATGATGAT 1019
QY 1642 TGTG-----TATCATCAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1693
Db 1020 TGTGTTTATCATATATCAATTAAGAAAGTATGATGATGATGATGATGATGATGATGATGATGAT 1079
QY 1694 ATATTTTCCACACTCATCAAAACC 1719
Db 1080 TTAATGTCTTCACTCTTCACACTC 1105
RESULT 3
CAPSY1 1295 bp mRNA linear PLN 05-FEB-1994
LOCUS CAPSY1
DEFINITION C.annuum psy1 mRNA for phytoene synthase.
ACCESSION X68017.1 GI:433993
VERSION X68017.1
KEYWORDS phytoene synthase.
SOURCE Capsicum annuum
ORGANISM Capsicum annuum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; Lamiales; Solanales; Solanaceae; Capsicum.
REFERENCE 1 (bases 1 to 1295)
Romer S., Huquaney P., Bouvier F., Camara B. and Kuntz M.
Expression of the genes encoding the early carotenoid biosynthetic
enzymes in Capsicum annuum
Biochem. Biophys. Res. Commun. 196 (3), 1414-1421 (1993)
JOURNAL MEDLINE 8250898
PUBMED 2 (bases 1 to 1295)
AUTHORS Kuntz M.
TITLE Direct Submision
REFERENCE Submitted (30-JUL-1992) M. Kuntz, Inst. de Biologie Molec. des
Plantes, 12 rue du Général Zimmer, 67084 Strasbourg, FRANCE
FEATURES
source
1..1295
/location/Qualifiers
1..1295
/organism="Capsicum annuum"
/mol_type="mRNA"
/strain="Lamuyo"
/db_xref="taxon:4072"
/dev_stage="ripening fruits"
1..1295
/gene="psy1"
1..1260
/gene="psy1"
/notes="subunit structure monomer"
/codon_start=1
```

```

/product="phytoene synthase"
/protein_id="CAA48155.1"
/db_xref="GI:433994"
/translation="MSVALLWVSEPCDVSNSTGFLVSVREGRNIFDSSGRNLACNER
IKRGGKRWSPGSLGAGQSGSKRSALVATPGEKMTSSERVAVIVRQAA
LVKQLRSTDELVDVKDIPITGLSLSEADRCSEVCAKTFYLDLMTBERK
AIWAIYWRRTDELVDGNASHITPDLDMEDLEBVFSGRPFDMLADLSTVSK
EPVDIOPRDMIEGRMDLRKSRVNFDELFLCYVAGTVGVPIMJGAPSKAT
TESYVNAALAIAGINOLNLIIRLDVGEDARGRVYLPODELQAGLSDSDIFAGVTTK
WRIEMKOIORARKEPDEARKEVTELSAASMPULASILLYRILDEIENDVNNFTK
RAVSKERKLLALPIATVAKSLVPSTRT"
BASE COUNT      379 a      219 c      361 g      336 t
ORIGIN
Query Match      49.5%; Score 856.2; DB 8; Length 1295;
Best Local Similarity 81.5%; Pred. No. 2.4e-198;
Matches 1057; Conservative 0; Mismatches 198; Indels 42; Gaps 4;
OY 250 ATGTCGTGCTTGTGTTGGGTTGTTCTCCCACTTCGAGTCTCGAATGGACAGGA 309
DB 1 ATGTCGTGCTTGTGTTGGGTTGTTCTTC---CTTGAGAGTCTCAAGCGGACAGGA 57
OY 310 TTGTTGATTCAGTCCGAGAAAGAAACCGCGCTTTGTTATCATCCAGTTCTTACGTCGA 369
DB 58 TTCTGTGATCCGTTCTGAGGAGAAACCGGATTTTGATTCGTCGGGG----- 105
OY 370 GATGGAATTTGATGGATGGGAGATCAAGA---AAGTGGAGACAAAGGTGAGT 426
DB 106 CGTAGAATTTGGCGTGAAGAGAAATCAAGAGAGAGGTGAGAAACAAAGGTGAGT 165
OY 427 TTTGGCTCTTAAATGTGATCCAGATATTCATCTTGCGTGATCAAGACTGAAAG 486
DB 166 TTTGG-----TTCTTACTTGGGAGAGCAAAACTGGAAGT 201
OY 487 GGAAAGACTTCTCTGTAAGTCCAGTTTGGTGGCTAGCCCACTGAGAAATGACTGTG 546
DB 202 GGAGCGAAATTTCTGTAAGTCTCTGTAAGTCTGCTAGCTAGCCGCTGAGAAATGACTGTG 261
OY 547 TCATCAGAGAAAGAGTATGATGTGATTTAAGACAGGACGTTTGTAGAGAGGAG 606
DB 262 TCATCAGAGAGAGTATGATGTGATTTTGAAGGACAGGACCTTGTGAGAGAGAG 321
OY 607 CTGAGATCTACCGATGATTTAGAAAGTGAAGCCGATTTGTTTCCAGGGAATTTGGGC 666
DB 322 CTGAGATGACCGAGATGATTTAGTGAAGAAATATACATTTCCGGGGACTTTGGGC 381
OY 667 TTGTTGAGTGAAGCATGATGCTGTTGGCGAAGTATGCAAGTATGCAAAAGCATTT 726
DB 382 TTGTTGAGTGAAGCATGATGATGCTGTTGGCGAAGTATGCAAGTATGCAAAAGCATTT 441
OY 727 TACTTAGAACCAAGCTATGACCCCGAGAGAGAAAGAGATTTGGGCAATATATGTTG 786
DB 442 TACTTAGAACCAAGCTATGATGCTCCGAGAGAGAAAGGCTATTTGGGCAATATATGTTG 501
OY 787 TGTGTGACAGAGAACGATGAGCTTGTGATGGCCCTTAATGATCCACATTAATCTCCGAA 846
DB 502 TGTGTGACAGAGAACGATGAGCTTGTGATGGCCCTTAATGATCCACATTAATCTCCGAA 561
OY 847 GCTTTAGATGAGTGGAGACCAAGCTGGAAGATATTTCAAGTGGCGGCCAATTGATATG 906
DB 562 GCTTTAGATGAGTGGAGACCAAGCTGGAAGATATTTCAAGTGGCGGCCAATTGATATG 621
OY 907 CTGTATGCTGCTTTATCCGATATGCTCCCAAGATTTCTGTTGATATTCAGCCATTCGA 966
DB 622 CTCGATGCTGCTTTGTCGACACAGTTTCCAAATTTCCAGTTGATTCAGCCATTCGA 681
OY 967 GATATGATTTGAAGAGATCGTATGAGACTTGTGGAATCCAGATACAAACTTTCGATAG 1026
DB 682 GATATGATTTGAAGAGATCGTATGAGACTTGTGGAAGATCAAGATACAAACTTTCGATAG 741
OY 1027 CTATATCTCTATTTGTTACTATGTTGCTGTGATCTGTAGAGATGATGAGTTCCTCAATTATG 1086

```

```

DB 742 CTATATCTCTATTTGTTATTTAGCTGTGATCGGTTGGTGTGATGAGTGTTCATATATG 801
OY 1087 GGATATGACCTGATATCAAGAGGCAACAGAGAGTATATATATGCTGCTTTGGCTTTA 1146
DB 802 GGATATGACCTGATATCAAGAGGCAACAGAGAGTATATATATGCTGCTTTGGCTTTG 861
OY 1147 GGGCTTGCMAATCAACTAACCAATATATCTAGAGATGTAGAGAGATGCGCAGAGAGGA 1206
DB 862 GGATATGACCAATCACTGACCAACATATCTTAGAGATGTGAGAGAGATGCGCAGAGAGGA 921
OY 1207 AGATATATCTGCTCTCAAGATGATTTAGCAGAGCAGGCTCTCCGCGAGAGATATTT 1266
DB 922 AGATATATCTGCTCTCAAGATGATTTAGCAGAGCAGGCTCTCTCCGAGAGATATTT 981
OY 1267 GCTGGAAGAGTACTGATTAATGTAGAGAACTTTATGAAAGAAACAAATTCAGAGGCGAGG 1326
DB 982 GCTGGAAGAGTACTGATTAATGTAGAGAACTTTATGAAAGAAACAAATTCAGAGGCGAGG 1041
OY 1327 AAATTTCTTTGATGATGAGAGAGAAAGTGTCAAGAACTGGACTTCTGTTAGTATGAGCTT 1386
DB 1042 AAGTTCTTTGACGAGGACAGAGAAAGAGTACCGAATTTGAGCGCAGCTAGTATGAGTGCCT 1101
OY 1387 GTGTTAAGAGCGCTGCTGTTGTTATCGCAATATTTGACAGAGATGGAAGCCAGACTAC 1446
DB 1102 GTGTTGACATCTGCTGTTGTTATCCGAGATATCTGACAGATGGAAGCCAGACTAC 1161
OY 1447 AACAACTTCAACAGAGAGGCTTATGTAGCAAGCCAAAGAGCTTCTCACCTTGCCTT 1506
DB 1162 AACAACTTCAACAGAGAGGCTTATGTAGCAAGCCAAAGAGCTTATGTAGCTTACTT 1221
OY 1507 GCTTATGCAAAATCTCTTGTGCCCCCTTAATGAACTT 1543
DB 1222 GCATATGCAAAATCTCTTGTGCTTCTTACAAAGACAT 1258

```

RESULT 4

TOMCBE

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

FEATURES

source

1. 1786

/organism="Lycopersicon esculentum"

/mol\_type="mRNA"

/db\_xref="taxon:4081"

/tissue\_type="fruit"

/dev\_stage="breaker fruit stage"

/tissue\_lib="Clontech fruit cDNA"

314..1552

/codon\_start=1

/product="phytoene synthetase"

/protein\_id="AAA4153.1"

/db\_xref="GI:170416"

/translation="MSVALLWVSPGVNSTGFLVSVREGRNIFDSSGRNLACNER

IKRGGKRWSPGSLGAGQSGSKRSALVATPGEKMTSSERVAVIVRQAA

LVKQLRSTDELVDVKDIPITGLSLSEADRCSEVCAKTFYLDLMTBERK

AIWAIYWRRTDELVDGNASHITPDLDMEDLEBVFSGRPFDMLADLSTVSK

EPVDIOPRDMIEGRMDLRKSRVNFDELFLCYVAGTVGVPIMJGAPSKAT

TESYVNAALAIAGINOLNLIIRLDVGEDARGRVYLPODELQAGLSDSDIFAGVTTK

WRIEMKOIORARKEPDEARKEVTELSAASMPULASILLYRILDEIENDVNNFTK

RAVSKERKLLALPIATVAKSLVPSTRT"

Original source text: Lycopersicon esculentum (library: Clontech fruit cDNA) breaker fruit stage fruit cDNA to mRNA.

Location/Qualifiers





BASE COUNT 529 a 249 c 388 g 480 t  
ORIGIN

Query Match 48.7%; Score 842; DB 6; Length 1646;  
Best Local Similarity 74.3%; Pred. No. 7e-195;

Matches 1222; Conservative 0; Mismatches 330; Indels 92; Gaps 9;

101 TATAATCGTTGAATTAAGTAGATAGACTAGTGTGATATCTAC---AAGTATTTGGTTT 157  
Db TAAATTTGTTGAGAGTGAATATTTCTAGTGGAACTCTAGAGATATTTATTTT 112  
158 TGATTAATAGGCTAGAGTGAAGATTAATTAAGAAAGACAAAACTTGGAAATGT 217  
Db CTATTAACCTAAGTAAGTTTGAAGGTGACA-AAAAGAAAGACAAAACTTGAATGT 171  
218 TTTAGACACGAGGTTTCTGTTTCATGAGATGCTGCTGTTTGTGTTGTTT 277  
Db TTTAGACACCAAGGTTTCTGTTTCATGAGATGCTGTTTGTGTTGTTT 228  
278 CTCCCACTTCCGAGGCTTCGAATGGAGACAGATTTGTTGATTCAGTCCGAGAAGAAAC 337  
Db CTC---CTTGTGAGCTCTCAATGGAGCAAGTTTCATGAAATCAAGTCCGGAGGGAACC 285  
338 GCGTCTTGTATCATCCAGGTTCTAGCTCGAGATGGAATTTGATGGAATGGAGAA 397  
Db GTTTTGTGATTCATCGAGG-----CATAGGAATTTGGTTCATGAGAGAA 333  
398 TCAAGAAAGGTGGAGACAAAGTGAATTTTGGCTTTATTTGCTGATCCAAATAT 457  
Db TCAATAGAGT----- 344  
458 CATGCTGGTGGTGCATCAAGAACTGAAGAGGACCTTCTCTGACAGTCCAGTTGG 517  
Db -----GTTGGAAGACAACTAATATGACGGAATTTTCTGACGCTCTGCTATTT 396  
518 TGGCTAGCCCAAGCTGGAAGAAATGACTGTCTCATGAGAAAGATGATGATGAT 577  
Db TGGTACTCTCATCTGAGAACGAGCATGACATCGGAACAGATGCTATGATGATGAT 456  
578 TAAAGCAGGAGCTTTAGAGAGAGGAGAGCTGAGATCTACGATGATTTAGAGGAGC 637  
Db TGAGCAGGAGGAGCTTTGAGAGAGGCACTGAGATCTACCAATGATTTAGAGGAGC 516  
638 CGGATATTGTTGTTCCAGGAAATTTGGGCTTGTGATGAGCAATGATGCTGTGAGC 697  
Db CGGATATACCTATTTCCGGGAATTTGGGCTTGTGATGAGCAATGATGATGATGAT 576  
698 AAGTATGTCAGAGATGCAAGACATTTTACTAGGAACCAAGCTAATGACCCAGAGA 757  
Db AAGTATGTCAGAGATGCAAGACATTTTACTAGGAACCAATGATGATGATGATGAT 636  
758 GAAGAAAGACTTCTGGGCAATATATGTGTGTCAGAGAAAGAGATGCTTTGATG 817  
Db GAAGAAAGGCTATCTGGGCAATATATGTGTGTCAGAGAAAGAGATGCTTTGATG 696  
818 GCCCTAATGATCCCAATCACTCCGCAAGCTTTGATGAGGAGACCAAGCTGAGAG 877  
Db GCCCAAGGCAATCAATATATTTACCCGCAAGCTTTGATGAGGAGACCAAGCTGAGAG 756  
878 AATATTTGAGTGGGCGCCATTTGATGATGCTGATGCTGCTTATCCGATCTCTCCA 937  
Db ATGTTTCAATGGGGCGCCATTTGATGATGCTGATGCTGCTTATCCGATCTCTCCA 816  
938 GATTTCTGTTATATTCAGCATTCAGAGATGATGATGAGAAATGCTATGAGACTGT 997  
Db ACTTTCAGATTGATATTCAGCATTCAGAGATGATGATGAGAAATGCTATGAGACTGT 876  
998 GGAATTCAGATACAAACCTTTCAGATGAGCTATATCTGATGATGATGATGATGAT 1057  
Db GGAATTCAGATACAAACCTTTCAGATGAGCTATATCTGATGATGATGATGATGAT 936  
1058 CTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1117

Db 937 CGTTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 996  
Qy 1118 AGAGTAT 1177  
Db 997 AGAGGAT 1056  
Qy 1178 GAGATGTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1237  
Db 1057 GAGATGTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1116  
Qy 1238 AGGAGGCTCTCCGACAGACATATTTGCTGAGAGTGAATGATGATGATGATGATGATGAT 1297  
Db 1117 AGGAGGCTCTCCGACAGACATATATTTGCTGAGAGGATGATGATGATGATGATGATGAT 1176  
Qy 1298 TTAGAAGAAACAAATATGAGAGGAGAGAAATTTGATGAGATGATGATGATGATGATGATGAT 1357  
Db 1177 TTAGAAGAAACAAATATGAGAGGAGAGAAATTTGATGAGAGGAGAGAGAGAGAGAGAG 1236  
Qy 1358 CAGAACTGAGTCTGCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1417  
Db 1237 CAGAACTGAGTCTGCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1296  
Qy 1418 TATTGACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1477  
Db 1297 TACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1356  
Qy 1478 AGGCAAG-AGGCTTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1536  
Db 1357 AATCAAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1416  
Qy 1537 AGAATCTCTCTCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1596  
Db 1417 AAGCTGCTCTCTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1476  
Qy 1597 ACT-----AAGAACTGAGTACTGTTAATGATGATGATGATGATGATGATGATGAT 1648  
Db 1477 AATGATACATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1536  
Qy 1649 CATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1704  
Db 1537 CATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1596  
Qy 1705 ACATCATCAACCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1728  
Db 1597 ACTTATCTCAACTCTTTGAGAG 1620

RESULT 7  
LRIPE 1614 bp mRNA linear PLN 11-MAY-1995  
DEFINITION Tomato fruit ripening specific mRNA.  
ACCESSION Y00521  
VERSION Y00521.1 GI:19340  
KEYWORDS unidentifed reading frame.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asterales; Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
REFERENCE 1  
Ray, J., Bird, C., Maunders, M., Grierson, D. and Schuch, W.  
Sequence of pTOM5, a ripening related cDNA from tomato  
Nucleic Acids Res. 15 (24), 10587 (1987)  
MEDLINE 88096591  
PUBMED 3697097  
REFERENCE 2 (bases 1 to 1614)  
AUTHORS Bird, C.R.  
TITLE Direct Submission  
JOURNAL Submitted (09-DEC-1987) Bird C.R., Imperial chemical industries  
PLC, Plant biotechnology group, Po Box 11, The Heath, Runcorn,  
Cheshire, WA7 4OE, U.K.  
FEATURES  
source  
1..1614

```

/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/strain="M11"
/cultivar="Alisa Craig"
/db_xref="taxon:4081"
/clone="p10M5"
201.1439
/note="tomato fruit ripening specific URF"
/codon_start=1
/protein_id="CA68575.1"
/db_xref="GI:19341"
/db_xref="SWISS-PROT:P08196"
/translation="MSVALLWVSPCDVSNSTSMESVREGNREFDSRRHNLVSNR
INRGGKQTNNGRKFSVSAILATPSGERTWSEOMYDVLLRQALVYKQLRSTNEL
EVKEDIPGNGILSLSEAYDRGECVCAATPNGLMTLPERRALVLYWCRT
DELVDGPASVYTPAALDRMENRLEVDNRPDMLDGLSDTVSNPVDIOPRDMT
EGMRMDKRSRYKNPDELYCYVAVAGVGMSPVINGIAIESKATSESYNALAAG
IANQUNILRDVGDARGRVYLPQDELQAGLSDDEDIFAEKRVTDKRIEMKQIHRA
RKFPEAEKGVTELSASRFPVMSLVLYRKILDEIEANDYNNTKRAYVSKSDVDC
ITVICIKISCASYKNASLQR"
misc_feature
246..254
/note="putative glycosylation site"
misc_feature
702..710
/note="putative glycosylation site"
misc_feature
1329..1337
/note="putative glycosylation site"
misc_feature
1419..1427
/note="putative glycosylation site"
BASE COUNT      515 a      241 c      388 g      470 t
ORIGIN
Query Match      48.0%; Score 829.8; DB 8; Length 1614;
Best Local Similarity 76.5%; Pred.No.6.7e-192;
Matches 1143; Conservative 0; Mismatches 272; Indels 80; Gaps 7;
QY 101 TATAATCGTTGAATTAAGTAGATAGCTTAAGTGAATCTAC--AAGTATGGTTTT 157
Db 53 TAAATTTGTTGAGAGTGAATATATCTTAGTGGAATCTACGAGTAAATTAATTTT 112
QY 158 TGATTAATAAGCTGAGTGAGAGTAAACATTAAGAGAAACAAAACCTTGGGAATGT 217
Db 113 CTATTAATAAGTAAAGTTTGAAAGTGACA-AAAGAAAGACAAAACCTTGAATGT 171
QY 218 TTTAGACCAAGAGGTTCTTGTTTCATGACATGTCTTGCTTTGTTGGTGGTGT 277
Db 172 TTTAGACCAAGAGT---TTTCTTGCTCAGATGTCTTGCTTGTATGGTGT 228
QY 278 CTCCCACTTCGAGGCTCGAATGGAGACAGATTGGATTCACTCCGAGAGAAAC 337
Db 229 CTC---CTTGAGCGTCAATGGGACAAAGTTTCATGAATCAAGTCGGAGAGGAAAC 285
QY 338 GCGCTTTGTATCAATCCAGGTTCTTCAAGTCAAGATAGAAATTTGATGTGAATGGAGAA 397
Db 286 GTTTTGTGATTCATCGAG-----CATAGGAATTTGGTGTCAATAGAGAA 333
QY 398 TCAAGAAAGTGGAGACAAAGTGAATTTTGCTTTAATGCTGATCCAGATAT 457
Db 334 TCATTAAGGT----- 344
QY 458 CATGCTGGGTGATCAAGAACTGAAGAGAGACATTTCTGTACAGTCCAGTTGG 517
Db 345 -----GTTGAAAGCAAACTAATATATGACGGAATTTTCTGTACGCTGTATTT 396
QY 518 TGGCTAGCCCACTGAGAAATGACTGTGTCATGAGAAAAAGGTGATATGATG 577
Db 397 TGGCTACTCTCATCTGGAACGAGATGACATCGAAACAGATGCTATATATGATG 456
QY 578 TAAAGCAGGAGCTTTAGTGAAGAGCAGCTGAGATCTAACGATATTAGAAGTGAAGC 637
Db 457 TGAAGCAGGAGCCTTGTGTGAAGAGCAGCTGAGATCTAACATAGATTTAGAAGTGAAGC 516
QY 638 CGGATATTTGTTTCCAGGAATTTGGCTTTGATGGAACATATGATCGTTGGCG 697

```

```

Db 517 CGGATATACCTATTCGGGGAATTTGGGCTTTGTAAGTGAACATATGATGATG 576
QY 698 AAGTATGTCAGAGTATGCAAAAGACATTTACTTAGGAACCAAGCTATATGACCCAGAGA 757
Db 577 AAGTATGTCAGAGTATGCAAAAGCGTTTAACCTTAGGAACCTATGCTATATGCTCCAGAGA 636
QY 758 GAAGAAAGCTATCTGGGCAATATATGTGTGTGTCAGAGAAACGATAGACTGTGTTATG 817
Db 637 GAAGAAAGGCTATCTGGGCAATATATGTATGTGTGTCAGAGAAACGATAGACTGTGATG 696
QY 818 GCCCTATGATCCCAATATCTCCGACCTTATGATAGTGGAGAACACAGGCTGTAAG 877
Db 697 GCCCAACGATCATATATTAACCCGACCTTATGATAGTGGAGAAATAGGCTTAAG 756
QY 878 ATATTTTCAGTGGCGCCCAATTTGATATGCTGATGCTGTTATCCGATACTCTCCA 937
Db 757 ATGTTTCAATGGCGCCCAATTTGATGATGCTGATGCTGATGCTGATGCTGATGCTGAT 816
QY 938 GATTTCTGTTGATATTCAGCCATTCAGAGATATGATGAAGAAATCGTATGACTTTGT 997
Db 817 ACTTTCAGTTGATATTCAGCCATTCAGAGATATGATGAAGAAATCGTATGACTTTGA 876
QY 998 GGAATTCAGATACAAAACCTTGTGATGATGATGATGATGATGATGATGATGATGATG 1057
Db 877 GAAATTCAGATACAAAACCTTGCACGACCTATACCTTTATATGTTATATGTTGCTGATA 936
QY 1058 CTGTAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1117
Db 937 CGGTTGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 996
QY 1118 AGAGTATATATATGCTGCTTTGCTTTAGGGCTTGGCAATCAATCAATCAATCAATCAAT 1177
Db 997 AGAGCTATATATATGCTGCTTTGCTTGGGATCGCAATCAATCAATCAATCAATCAATCAAT 1056
QY 1178 GAGATGTAAGAGAAATGCCAGAAAGAGAGATGATCTTGCTTCAAGATGATGATGATGATG 1237
Db 1057 GAGATGTTGAGAGAAATGCCAGAAAGAGATGATCTTGCTTCAAGATGATGATGATGATG 1116
QY 1238 AGGAGAGGCTCTCCGAGAAAGACATATTTGCTGGAAGATGATGATGATGATGATGATG 1297
Db 1117 AGGAGAGGCTCTATCCGAGAAAGATATATTTGCTGGAAGGATGATGATGATGATGATGATG 1176
QY 1298 TTAGAGAAACAAATTCAGAGGGCGAGAAATTTCTTTGATGATGATGATGATGATGATG 1357
Db 1177 TTAGAGAAACAAATTCATAGGCGAGAAAGTTCTTTGATGATGATGATGATGATGATGATG 1236
QY 1358 CAGAACTGAGACTGCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1417
Db 1237 CAGAAATGAGCTGCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1296
QY 1418 TATTGACGAGATGAGCCACGACTACAAACATTCACAAAGAGGGCTTATGATGATGATG 1477
Db 1297 TACTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1356
QY 1478 AGCCAAAG-AAAGCTTCACTGCTGCCATTCCTTATGCAAAATCTCTTGCCCTTAT 1536
Db 1357 AATCAAGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1416
QY 1537 AGAATCTCTCTCCACTACGAAAGACATGATGAATGATGATGATGATGATGATGATGATG 1591
Db 1417 AAAATGCTCTCTTCAAAAGATTAAGCATGAATGAATGATATATATATATATATAT 1471

```

RESULT 8  
 A68204 LOCUS 1239 bp DNA linear PAT 06-MAY-1999  
 DEFINITION Sequence 2 from Patent WO9746690.  
 ACCESSION A68204  
 VERSION A68204.1 GI:4759372  
 KEYWORDS  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 1239)  
 Drake, C.R., Bird, C.R., and Schuch, W.W.  
 TITLE ENHANCEMENT OF GENE EXPRESSION  
 JOURNAL Patent: WO 9746690-A 2 11-DEC-1997;  
 ZENECA LTD (GB)

FEATURES  
 source Location/Qualifiers

1..1239  
 /organism="Lycopersicon esculentum"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4081"  
 /clone="GTOM5 - PHYTOENE SYNTHASE GENE"

BASE COUNT 377 a 207 c 317 g 338 t  
 ORIGIN

Query Match 46.9%; Score 811.2; DB 6; Length 1239;  
 Best Local Similarity 83.8%; Pred. No. 2.4e-187;  
 Matches 917; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 466 GGTGGATCAAGAACTGAAAAGGAGCACTTCTCTGTACAGTCCAGTTGGTGCTAGC 525  
 DB 145 GTTGAAAGCAAACTAATATGACGGAATTTTCTGACGCTCTGCTATTTGGCTACT 204  
 QY 526 CCAGCTGAGAAATGACTGTGTATCAAGAAAAGGTGTATGATGTGTATTAAGCAG 585  
 DB 205 CCATCTGAGAAACGAGCATGACATCGAACAGATGTCTATGATGTGTGTGAGGCGAG 264  
 QY 586 GCAGCTTAAAGCAAGGCAAGTGTATACCGATGATTTGAAGTGAAGCCGAGATTT 645  
 DB 265 GCAGCTTGTGAAGGCAAGTGTATACCGATGATTTGAAGTGAAGCCGAGATTTA 324  
 QY 646 GTTGTTCAGGGAATTTGGGCTTTGTGATGAAGATATGATGCTTTGGCCAGATATGT 705  
 DB 325 CCTATTCGGGGAATTTGGGCTTTGTGATGAAGATATGATGCTTTGGCCAGATATGT 384  
 QY 706 GCAGATATGCAAGACATTTTACTTAAGAACCAAGCTATATGACCCCAAGAGAAAGA 765  
 DB 385 GCAGATATGCAAGACATTTTACTTAAGAACCAAGCTATATGACCCCAAGAGAAAG 444  
 QY 766 GCTATCTGGGCAATATATGTGTGTGCAAGAAACGATGATGCTTTGGCCCTATAT 825  
 DB 445 GCTATCTGGGCAATATATGTGTGTGCAAGAAACGATGATGCTTTGGCCCTATAT 504  
 QY 826 GCATCCCAATTAATCCCGCAAGCTTTAGTAGTGGAGACCGAGCTGGAGATATTTTC 885  
 DB 505 GCATCAATATTAATCCCGCAAGCTTTAGTAGTGGAGAAATGGCTTGAGAGATTTTC 564  
 QY 886 AGTGGCGGCAATTTGATATGCTGTGATGCTTTATCCGATATGCTCCAGATTTCT 945  
 DB 565 AATGGCGGCAATTTGATATGCTGTGATGCTTTATCCGATATGCTTTCACTTTCA 624  
 QY 946 GTTGTATTCAGCCATTCAGAGATATGATTTGAAGAAATGCGATGCTTTGGAAATCC 1005  
 DB 625 GTTGTATTCAGCCATTCAGAGATATGATTTGAAGAAATGCGATGCTTTGGAAATCC 684  
 QY 1006 AGATCAAAACCTTGGATGAGTATATCTATTTTACTATTTGCTGTGTGCTGTAGTA 1065  
 DB 685 AGATCAAAACCTTGGATGAGTATATCTTTATTTTATTTATGTGTGTGTGCTGTAGTA 744  
 QY 1066 TTGATGAGTGTTCAGTATGAGTATGCACTGTGATCAAAAGCAAAACAGAGAGTGA 1125  
 DB 745 TTGATGAGTGTTCAGTATGAGTATGCACTGTGATCAAAAGCAAAACAGAGAGTGA 804  
 QY 1126 TATATGCTGCTTTGGCTTTAGGGCTTTGCAATCACTTACCAATATATCTCAGAGATGA 1185  
 DB 805 TATATGCTGCTTTGGCTTTAGGGCTTTGCAATCACTTACCAATATATCTCAGAGATGA 864  
 QY 1186 GGAAGAGATGCGAAGAGAGAGATATCTTGGCTCAAGATGATTTGGCAAGGCAAGG 1245  
 DB 865 GGAAGAGATGCGAAGAGAGAGATATCTTGGCTCAAGATGATTTGGCAAGGCAAGG 924  
 QY 1246 CTCTCCGAGAGACATATTTGCTGGAAGAGTGAAGTGAAGTGAAGCACTTTATGAAG 1305

DB 925 CTATCCGATGAGATATATTTTGTGGAAGGGGTGACCCGTAATGAGAAATCTTTATGAAG 984  
 QY 1306 AACCAATTCAGAGGGGAGGAGAAATTTCTTGATGAGTCAAGAAAGGTGTCACAGACTG 1365  
 DB 985 AACCAATTCAGAGGGGAGGAGAAATTTCTTGATGAGTCAAGAAAGGTGTCACAGACTG 1044  
 QY 1366 GACTCTGTAGTATGAGTGGCTGTGTATCAAGCGCTGTCTGTGTATCGCAAGATATGAGC 1425  
 DB 1045 AGCTGAGTATGATGATTCCTGTATGGGCAATCTTTGTGCTGTATCCGCAAAATACTAGT 1104  
 QY 1426 GAGATTAAGCCAAAGCACTTACCAACATTCACAAAGAGGGCTTATGTAGCAAGCCAAAG 1485  
 DB 1105 GAGATTAAGCCAAAGCACTTACCAACATTCACAAAGAGGGCTTATGTAGCAAGCCAAAG 1164  
 QY 1486 AAGCTTCCAGCTTGGCCCATGCTTATGCAAAATCTCTTGGCCCTTAATGACCTCC 1545  
 DB 1165 AAGTGAATGCAATTAATGCAATGCAATGCAAAATCTCTTGGCCCTTAATGACCTCC 1224  
 QY 1546 TCTCCACTAGCAAA 1559  
 DB 1225 TCTCTCAAGATA 1238

RESULT 9  
 AR156064 1239 bp DNA linear PAT 08-AUG-2001  
 LOCUS  
 DEFINITION Sequence 2 from patent US 6239331.  
 ACCESSION AR156064  
 VERSION AR156064.1 GI:15124117  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 Unclassified.

REFERENCE  
 1 (bases 1 to 1239)  
 Drake, C.R., Bird, C.R., and Schuch, W.W., Walter.  
 TITLE Enhancement of tomato phytoene synthase gene expression with a modified DNA

JOURNAL Patent: US 6239331-A 2 29-MAY-2001;

FEATURES  
 source Location/Qualifiers

1..1239  
 /organism="unknown"  
 BASE COUNT 377 a 207 c 317 g 338 t  
 ORIGIN

Query Match 46.9%; Score 811.2; DB 6; Length 1239;  
 Best Local Similarity 83.8%; Pred. No. 2.4e-187;  
 Matches 917; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 466 GGTGGATCAAGAACTGAAAAGGAGCACTTCTCTGTACAGTCCAGTTGGTGCTAGC 525  
 DB 145 GTTGAAAGCAAACTAATATGACGGAATTTTCTGACGCTCTGCTATTTTGGCTACT 204  
 QY 526 CCAGCTGAGAAATGACTGTGTATCAAGAAAAGGTGTATGATGTGTATTAAGCAG 585  
 DB 205 CCATCTGAGAAACGAGCATGACATCGAACAGATGTCTATGATGTGTGTGAGGCGAG 264  
 QY 586 GCAGCTTAAAGCAAGGCAAGTGTATACCGATGATTTGAAGTGAAGCCGAGATTT 645  
 DB 265 GCAGCTTGTGAAGGCAAGTGTATACCGATGATTTGAAGTGAAGCCGAGATTTA 324  
 QY 646 GTTGTTCAGGGAATTTGGGCTTTGTGATGAAGATATGATGCTTTGGCCAGATATGT 705  
 DB 325 CCTATTCGGGGAATTTGGGCTTTGTGATGAAGATATGATGCTTTGGCCAGATATGT 384  
 QY 706 GCAGATATGCAAGACATTTTACTTAAGAACCAAGCTATATGACCCCAAGAGAAAGA 765  
 DB 385 GCAGATATGCAAGACATTTTACTTAAGAACCAAGCTATATGACCCCAAGAGAAAG 444  
 QY 766 GCTATCTGGGCAATATATGTGTGTGCAAGAAACGATGATGCTTTGGTGTAGGCCCTATAT 825  
 DB 445 GCTATCTGGGCAATATATGTGTGTGCAAGAAACGATGATGCTTTGGTGTAGGCCCAATC 504



OY		826	GCATCCGCAATAACCTCGCAAGCTTTTAAATAGTGGGAGACACAGGCTGGGAAGATTATTC	895
Db		505	GCACTATATTAATCCCCGCAAGCCTTAAATAGTGGGAAAATAGGCTAAGAAGATGTTC	564
OY		886	AGTGGCGCGCCATTGTAATATGCTTGANTCTGCTTATTCGATACATGCTGCCAATTTCT	945
Db		565	AATGGCGCGCATTTTGAATGCTGATGCTGATGCTGTTTTGTCCGATACAGTTTCTAACCTTCCA	624
OY		946	GTTGATATTACAGCCATTCAAGATATGATTTGAAGAAATGCGTATATGACTTTGGAAATCC	1005
Db		625	GTTGATATTACAGCCATTCAAGATATGATTTGAAGAAATGCGTATATGACTTTGAAGAAATCC	684
OY		1006	AGATPACAACCTTTGATGAGCTAATATCTAATTTGTCCTAATGTTGCTGGTATCTGTGGA	1065
Db		685	AGATPACAACCTTTGATGAGCTAATATCTAATTTGTCCTAATGTTGCTGGTATCTGTGGA	744
OY		1066	TTCATGATGTTTCCAGTATATGGGTATTCACCTGTAATCAAAGGCAACAAGAGAGTGT	1125
Db		745	TTGATGATGTTTCCAGTATATGGGTATTCACCTGTAATCAAAGGCAACAAGAGAGTGT	804
OY		1126	TATAATGCTGCTTTGGCTTTTAGGGCTTGCAATCACTAACCAATATATCTOAGAGATGA	1185
Db		805	TATAATGCTGCTTTGGCTTTTAGGGCTTGCAATCACTAACCAATATATCTOAGAGATGT	864
OY		1186	GGAGAAATGTCAGAAAGAGAAAGACTATATCTTGCTCAAGATGAAATTTAGCACAGGAGG	1245
Db		865	GGAGAAATGTCAGAAAGAGAAAGACTATATCTTGCTCAAGATGAAATTTAGCACAGGAGGT	924
OY		1246	CTCTCCGACGAAGACATATTTTCTGGAAGAGTGACTGATTAAGTGGAGGAACTTTATGAA	1305
Db		925	CTATCCGATGAAGATATATTTTCTGGAAGAGGAGTGAATTAAGAGAAATCTTTATGAA	984
OY		1306	AAACAATTCAGAGGGCCAGAGAAATCTTTGATGATCAGAGAAAGTGTCAACAGACTG	1365
Db		985	AAACAATTCAGAGGGCCAGAGAAAGTCTTTGATGATCAGAGAAAGGCGTACAGAAATG	1044
OY		1366	GACTCTGTAGTATGATGATGCGCTGTGTTAACAGGCTGCTGTTGATACGAATATTTGAC	1425
Db		1045	AGCTGAGCTAGTATGATGATGCGCTGTGTTGATGATGCGGAAAAATATCTAAT	1104
OY		1426	GAGATTGAAGCCAGACTACCAACAATCTACAGAGAGGCTTATGTTAGCAAGCCAAAG	1485
Db		1105	GAGATTGAAGCCAGACTACCAACAATCTACAGAGAGGCAATATGATGCAAAATCAAAG	1164
OY		1486	AAGCTTCTCACCTTGCCCATGCTTATGCAAAATCTTTGTCGCCCTAATAGAACTTCC	1545
Db		1165	AAGTGAATGCACTTACATATGCAATGCAAAATCTTTGTCGCCCTAATAGAAATCTCC	1224
OY		1546	TCTCCACTAGCAAA 1559	
Db		1225	TCTCTTCAAGATA 1238	

RESULT 10  
 BD005486  
 LOCUS  
 DEFINITION Enhancement of gene expression.  
 ACCESSION BD005486  
 VERSION BD005486.1 GI:18633857  
 KEYWORDS JP 2001501810-A/2.  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 1239)  
 Drake,C.R., Bird,C.R. and Schuch,W.W.  
 Enhancement of gene expression  
 Patent: JP 2001501810-A 2 13-FEB-2001;  
 ZENECA LTD  
 COMMENT OS Lycopersicon esculentum (tomato)  
 PN JP 2001501810-A/2  
 PD 13-FEB-2001

	PF	23-MAY-1997 JP 1998500302
	PR	07-JUN-1996 GB 9611981.3
	PI	CAROLINE RACHEL DRAKE, COLIN ROGER BIRD, WOLFGANG WALTER SCHUCH
	PC	C1ZNI5/67, C1ZNI5/82, C1ZNI5/29, C07K14/415
	CC	Strandedness: Double;
	CC	Topology: Linear;
	PH	Key
	FT	Location/Qualifiers
	FT	source
FEATURES		
source	Location/Qualifiers	
	1. 1239	
	/organism="Lycopersicon esculentum"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:4081"	
BASE COUNT	377 a 207 c 317 g 338 t	
ORIGIN		
Query Match	46.9%;	Score 811.2; DB 6; Length 1239;
Best Local Similarity	83.8%;	Pred. No. 2.4e-187;
Matches 917, Conservative	0;	Mismatches 177; Indels 0; Gaps 0;
OY	466	GGTGGATCAAGAACTGAGAAAAGGAAAGCACTTCTCTGTACAGTCCAGTTGGTGGCTAGC 525
DB	145	GGTGGAAAGCAAACTAATAATGACGGAAATTTCTGTACGGTCTCTATTTTGGCTACT 204
OY	526	CCAGCTGGAGAAATGACTGTGTCACTCAGAGAAAAGGTATGATGTGATTTAAAGCAG 585
DB	205	CCAATCTGGAGAACGACGATGACATCCGGAACAGATGTGTATGATGTGTTTGAAGCAG 264
OY	586	GCAGCTTTAGTGAAGAGGACAGCTGAGATCTACCGATGATTTTGAAGTGAACCCGGAATTT 645
DB	265	GCACCTCTGGGAGAGAGGCAACTGAGATCTACCAATGATGTTAGAGAAAGTGAACCCGGAATTA 324
OY	646	GTTGTTCCAGGGAATTTGGGCTTGTGTAACACATATGATCGTTGGGCGAAGTATGT 705
OY	325	CTTATCTCCGGGAAATTTGGGCTTGTGTGTAACATATGATAGTGTGTGAAAGTATGT 384
DB	706	GCAGAGTATGAAAGACATTTTACTTGAACCAAGCTAATGACCCGAGAGAGAAGA 765
DB	385	GCAGAGTATGAAAGACATTTTACTTGAACCAAGCTAATGACCTCCGAGAGAAGAAG 444
OY	766	GCTATCTGGGCAATATATGTGTGTGTCAGAGAAACGATAGCTTTGTTGATGTGCCCTAT 825
DB	445	GCTATCTGGGCAATATATGTGTGTGTCAGAGAAACAGATGAACTTTGTTGATGTGCCCAAC 504
OY	826	GCATCCCAACATTAATCCGCAAGCTTTTGAATGATGGGGAGACCAAGCTGGAAGATATTTTC 885
DB	505	GCATCAATATATTAATCCCGGACAGCTTGTATGTGTGGGAAATATAGCTTGAAGATGTTTTTC 564
OY	886	AGTGGGGGGCCATTTGATATGCTTGTATGCTGCTTTATCCGATGCTGTCCAGATTTTCT 945
DB	565	AATGGGGGGCCATTTGATATGCTGCTGATGTGCTTGTCCGATACAGTTTCTTAACCTTTTCCA 624
OY	946	GTTGATATTTCAAGCAATTCAGAGATATGATTTGAAGAAATGCTATGAGACTTGTGAAATCC 1005
DB	625	GTTATATTTTCAGCAATTCAGAGATATGATTTGAAGAAATGCTATGAGACTTGTGAAGAAATCC 684
OY	1006	AGATACAAAACCTTTGAGATGAGCTATATCTATTTGTTACTATGTTTGCTGTAGTCTGTAGGA 1065
DB	685	AGATACAAAACCTTTGAGATGAGCTATATCTATTTGTTACTATGTTTGCTGTAGTCTGTAGGA 744
OY	1066	TTGATGAGGTGTCAGATTATGGGTATTTGACCTGATATCAAGAGCAACACAGAGAGTGT 1125
DB	745	TTGATGAGGTGTCAGATTATGGGTATTTGACCTGATATCAAGAGCAACACAGAGAGCTGT 804
OY	1126	TATATATGCTGTTTGGCTTTTATGGGCTTGCAATCAATCAATATCTCAGAGATGTA 1185
DB	805	TATATATGCTGTTTGGCTTTTATGGGCTTGCAATCAATCAATATCTCAGAGATGTT 864
OY	1186	GGAAAGATGTCAGAAAGAGAAAGTATATCTTGCTCAAGATGATTTAGCAAGGCAAGG 1245
DB	865	GGAAAGATGTCAGAAAGAGAAAGTATATCTTGCTCAAGATGATTTAGCAAGGCAAGGT 924



QY	1246	CTCTCCGACGAACATATTGTCCTGGAAGGTGATGATTAAGTGAAGAACTTTATGAG	1305		
Db	925	CTATCCGATGAAGATATATTGTCGGAAGGTGACCCGATTAATGAGAAATCTTTATGAG	984		
QY	1306	AAACAAATTCAGAGGGCGAGGAAATTCCTTTGATGATGATGAGAAAGGTGTCAAGAACTG	1365		
Db	985	AAACAAATATACATAGGGCAAGAAAGTTCTTTGATGAGGAGAAAGGCGTGACGAATTG	1044		
QY	1366	GACTCTGTAGTAAATGGCTGTGTGTTAAACAGCGCTGCTGTTGTATCGCAAGATATTGAC	1425		
Db	1045	AGCTCAGGTAGTAGTATCCCTGTATGGGCATCTTTGGTCTTGTAACCGAAATATCTAGAT	1104		
QY	1426	GAGATTGAAGCAACGACTACAAACAATCTCAAGAGAGGCTTTATGTTAGCAAGCCAAG	1485		
Db	1105	GAGATTGAAGCAACGACTACAAACAATCTCAAGAGAGATATGTTAGCAAAATCAAG	1164		
QY	1486	AAGCTTCTACACCTTGCCCATGCTTATGCAAAATCTCTTGAGCCCTTAATAGAACTCC	1545		
Db	1165	AAGTTGATTGCAATTACCTATTGCAATATGCAAAATCTCTTGAGCCCTTAACAAACTGCC	1224		
QY	1546	TCTCCACTAGCAAA	1559		
Db	1225	TCTCTTCAAGATA	1238		
RESULT 11	LERYGTOMS	1355 bp	mRNA	linear	PLN 02-AUG-1993
LOCUS	1.esculentum (ry mutant)	GROM5 mRNA for mutant phytoene synthase.			
DEFINITION	X67144.1	GI:19346			
ACCESSION	GROM5 gene; mutant; phytoene synthase.				
KEYWORDS	Lycopersicon esculentum (tomato)				
SOURCE	Lycopersicon esculentum				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae, lamids; Solanales; Solanaceae; Solanum; Lycopersicon.				
REFERENCE	1 (bases 1 to 1355)				
AUTHORS	Fray, R.G. and Grierson, D.				
TITLE	Identification and genetic analysis of normal and mutant phytoene synthase genes of tomato by sequencing, complementation and co-suppression				
JOURNAL	Plant Mol. Biol. 22 (4), 589-602 (1993)				
MEDLINE	93344508				
PUBMED	8343597				
REFERENCE	2 (bases 1 to 1355)				
AUTHORS	Fray, R.G.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-JUL-1992) R.G. Fray, University of Nottingham, School of Agriculture, Sutton Bonington, Loughborough, Leicestershire LE12 5RD, UK				
FEATURES	source				
	1. .1355	Location/Qualifiers			
	/organism="Lycopersicon esculentum"				
	/mol_type="mRNA"				
	/strain="Alisa Craig"				
	/isolate="ry mutant"				
	/db_xref="taxon:4081"				
	/chromosome="3"				
	/clone="p5ty"				
	/tissue_type="pericarp"				
	/dev_stage="ripe fruit"				
	1. .1355				
	/gene="GROM5"				
	9. .1307				
	/gene="GROM5"				
	/codon_start=1				
	/product="mutant phytoene synthase"				
	/protein_id="CAA47625.1"				
	/db_xref="GI:19347"				
	/db_xref="SPTREMBL:O40166"				
	/translation="MSVALLVVSECDVSNQTSFMESVREGNRRPFDSRRHNLVSNERR				
	INRGKQKINNGKFSVRSALATPDSGERTWSQWYDVLRQALVQRJLRSTNEL				

BASE COUNT	412 a	223 c	354 g	366 t
ORIGIN				
Query Match	44.7%	Score 772;	DB 8;	Length 1355;
Best Local Similarity	84.5%;	Pred. NO. 9e-178;		
Matches 866;	Conservative 0;	Mismatches 159;	Indels 0;	Gaps 0;
variation /gene="GTOM5"				
1171..1335				
EAVPDPIDPGENILSLAAYRQCEVAEYAKETNLGMLTGMTPERRRAIYAIYWCPR				
DELVDYFNASTYITPAPALDRENTLEDYFVNGRPPEDMDGLSMTVSFNPVIOQPRDIT				
EGKRMIDLRKSRVYNFPELYLYCYVAGVATLMSVPMGLAIPEKATTESYNVALAAG				
IAQIQLTILIDVEDEDARGRVYLPQDELAAQASLVIDPFGVRYTKDIRIMPKQIHRA				
RKPFDEAEKVEITELISASRPVMAISLYAKRIIDLEIANDYNNFTKAAVYSKSNMLKD				
FFSFFSKSRGSNAITTVLGLACETFYIALIDRGPIDGTFP				
OY	466	GGTGGATCAAGAACTGAAAAGGAGAACACTTTCCTGTACAGTCACAGTTTGATGGCTAGC	525	
Db	153	GGTGGAAAGCAACTAATAATGACGGGAATTTTCTGTACGGCTGTGTAATTTGGCTACT	212	
OY	526	CCAGCTGGAGAAATGACTGTGTCAATCAGAAAAAGGTATGATGTGTATTAAGCAG	585	
Db	213	CCATCTGGAGAACGGAGATGACATCCGAAACAGATGCTATGATGTGTGTTTGAGCAG	272	
OY	586	GCAGCTTTAGTGAAGAGCGACGTGAGATCTAACGATATTAAGAAAGAACCGGATATT	645	
Db	273	GCAGCCTTTGGTGAAGAGGCAACTGAGATCTAACATAGTTAAGAAAGAACCGGATATA	332	
OY	646	GTTTGTTCACAGGAATTTGGGCTTTGTAGTGAAGCATATGATCGTTGTGGCGAAGTATG	705	
Db	333	CCATTTCCGGGGAATTTGGGCTTTGTAGTGAAGCATATGATGATGTGTGTGGAAGTATGT	392	
OY	706	GCAGAGTATCAAGACATTTTACTTAGAACCAGGTAATGACCCGAGAGAAAGAA	765	
Db	393	GCAGAGTATCAAGACATTTTACTTAGAACTATGCTAATGACTCCGAGAGAAAGAG	452	
OY	766	GCATATCTGGGCAATATATGTGTGTGTGTCAGAGAAACGATGACTTTGTGATGGCCCTAAT	825	
Db	453	GCATATCTGGGCAATATATGTATGTGTGTCAGAGAAACATATGAACTTTGTGATGGCCCAAC	512	
OY	826	GCATCCCAACATTAATCTCCGCAAGCTTTGTAGTAGGTGGGAGACCAAGGCTGGAAGATATTTTC	885	
Db	513	GCATCAATATATTAATCTCCGCAAGCTTTGTAGTAGGTGGGAAAAATAGGCTAAGAAATGTTTTTC	572	
OY	886	AGTGGGGGCGCATTTGATATGCTGTGATGCTGCTTATCCGATACCTGTCTCCAGATTTTCT	945	
Db	573	AATGGGGGCGCATTTGATATGCTGTGATGCTGCTTATCCGATACCTGTCTCCAGATTTTCCA	632	
OY	946	GTTGATATTTCAAGCATTTCAAGATATGATTAAGAAAGATGCGTATGACCTGTGGAATCC	1005	
Db	633	GTTGATATTTCAAGCATTTCAAGATATGATTAAGAAAGATGCGTATGACCTGTGGAAGAAATCG	692	
OY	1006	AGATACAAAACTTTGCATGAGCTATATCTGTAATTTGTACTATGTTGCTGTGTAACGTATGA	1065	
Db	693	AGATACAAAACTTTGCATGAGCTATATCTGTAATTTGTACTATGTTGCTGTGTAACGTATGA	752	
OY	1066	TTGATGAGTGTCCAGTTATGAGTATGCACTGAATCAAAAGGCAACACAGAGAGTGTA	1125	
Db	753	TTGATGAGTGTCCAAATTAATGGGATGCGCCCTGGAATCAAAAGGCAACACAGAGAGTGTA	812	
OY	1126	TATATATCTGCTTTGGCTTTTAGGGCTTGCAAAATCACTAACCAATATATCTGAGATGTA	1185	
Db	813	TATATATCTGCTTTGGCTTTTAGGGATGCAAAATCACTAACCAATATATCTGAGATGTT	872	
OY	1186	GGAGAAATGTCGAGAAAGAGAGATATCTGTCCTCAAAATGAAATTAGACACAGGCGAGG	1245	
Db	873	GGAGAAATGTCGAGAAAGAGAGATATCTTACTGCTCAAGATGATTAAGACACAGGCGAGGT	932	
OY	1246	CTCTCCAGCAAGACATATTTGTCTGGAAGAGTGAATGATTAAGTGAAGAACTTTATGAAG	1305	
Db	933	CTATCCGATGAAGATATATTTGTCTGGAAGAGGTGACGATTAATGGAAGAACTTTATGAAG	992	
OY	1306	AAACAAATTCAGAGGGCGAGGAATTTCTTTGATGAGTCAGAGAAAGGTGTCAAGAACTG	1365	

```

Db      993  AAACAAATACATAGGCGAAGAAATCTTTGATGAGGAGAGAAAGCGGTGACAGAAATTG 1052
Oy      1366  GACTGTGCTAGTAGATGCGCTGTGTAAACAGCGCTGCTGTGTATGCGAAGATATTGGAC 1425
Db      1053  AGCTAGCTAGTAGATGCTGTGTATGCGCATTTTGTCTGTATACCGCAAAATTAATTAAT 1112
Oy      1426  GAGATTGGAAGCCAGCACTCAACAACCTTCACAGAGAGCGCTTATGTTAGCAAGCCAAAG 1485
Db      1113  GAGATTGGAAGCCAGCACTCAACAACCTTCACAGAGAGCATATGTGAGCAAAATCAAT 1172
Oy      1486  AAGCT 1490
Db      1173  ATGCT 1177

RESULT 12
LOCUS   AF220218 1506 bp mRNA linear PLN 11-FEB-2000
DEFINITION Citrus unshiu phytoene synthase (Psy1) mRNA, complete cds.
ACCESSION AF220218
VERSION   AF220218.1 GI:6959859
KEYWORDS
SOURCE   Citrus unshiu
          Citrus unshiu
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosid II; Sapindales; Rutaceae; Citrus.
REFERENCE 1 (bases 1 to 1506)
            Kim,I.-J., Ko,K.-C., Kim,C.-S. and Chung,W.-I.
            Isolation of a cDNA encoding phytoene synthase from Citrus
            Unpublished
JOURNAL  2 (bases 1 to 1506)
            Kim,I.-J., Ko,K.-C., Kim,C.-S. and Chung,W.-I.
            Direct Submission
            Submitted (30-DEC-1999) Biological Sciences, Korea Advanced
            Institute of Science and Technology, 373-1 Kusong-dong, Yuseong-gu,
            Taejeon 305-701, South Korea
FEATURES
         source
         1..1506
            /organism="Citrus unshiu"
            /mol_type="mRNA"
            /cultiivar="Miyagawa (early)"
            /db_xref="taxon:55188"
            /cissue_type="fruit"
            /note="authority: Citrus unshiu (Swingle) Marc."
         gene
         1..1506
            /gene="Psy1"
            /gene="Psy1"
            /gene="Psy1"
            /codon_start=1
            /product="phytoene synthase"
            /protein_id="AAF3237.1"
            /db_xref="GI:6959860"
            /translation="MSVTLWVSPNSQLNCFGVDSVREENRLFYSSRFLYQHQR
            TAVNRSRKFQNNKCRNRSYPLDTDLRHPCSSGIDLPEISCMWAGCAVAMSSSE
            MVMVNVLKQALVNKQPSGTRDLVDNDILPGLSLLSAYDRSCGCAEVAKTFY
            LGTILMSEERRAIWAIYVRCRDELVDGNSAHSITPTALDRRESRLIEDLFRPDP
            MLDALSBTVTKRPVDIOPFRDMLRGMMDLRKSKYKRPDELILYLCYVATVGLMSV
            FVMGIALPDSQATTESVNAALALGIANOLTNILRDVGSDARGRYVLPQDELADAGLS
            DDDIFAGEVITIKRMFMNQIKRAMFEDMAENGTLELSEASRPVMASSLLLYRQILD
            EIEANDYNNFTRKRAVSKAKRIALPIVAYKSLRPSRIYLSKA"
         CDS
         1..1422
            /gene="Psy1"
            /product="phytoene synthase"
            /protein_id="AAF3237.1"
            /db_xref="GI:6959860"
            /translation="MSVTLWVSPNSQLNCFGVDSVREENRLFYSSRFLYQHQR
            TAVNRSRKFQNNKCRNRSYPLDTDLRHPCSSGIDLPEISCMWAGCAVAMSSSE
            MVMVNVLKQALVNKQPSGTRDLVDNDILPGLSLLSAYDRSCGCAEVAKTFY
            LGTILMSEERRAIWAIYVRCRDELVDGNSAHSITPTALDRRESRLIEDLFRPDP
            MLDALSBTVTKRPVDIOPFRDMLRGMMDLRKSKYKRPDELILYLCYVATVGLMSV
            FVMGIALPDSQATTESVNAALALGIANOLTNILRDVGSDARGRYVLPQDELADAGLS
            DDDIFAGEVITIKRMFMNQIKRAMFEDMAENGTLELSEASRPVMASSLLLYRQILD
            EIEANDYNNFTRKRAVSKAKRIALPIVAYKSLRPSRIYLSKA"
         BASE COUNT
         448 a 271 c 360 g 427 t
         ORIGIN
         Query Match 40.3%; Score 696.8; DB 8; Length 1506;
         Best Local Similarity 71.2%; Pred. No. 2,1e-159;
         Matches 949; Conservative 0; Mismatches 377; Indels 6; Gaps 2;
Oy      212  AATGTTTAAAGCCAGGAGTTTCTTTTCATAGAGATGCTGCTTTGTTGGG 271
Db      74  AATTTTCTTTTATCCAAAATAATATTTTACAGACATGCTGTTACATTGCTGGG 133
Oy      272  TTGTTTCCCACTTCGAGGCTCGAATGGGACAGATTGTTGATTCACTCGAGAAG 331

```

```

Db      134  TTGTATCACTTAATCAAAATTTGCCAATTTGCTTCGGGTTCTGTGATTCAGTTGAGAGG 193
Oy      332  GAAACCCGCTTTTGTATCAATCCAGTGTCTCACTGATGAGATTTGATGAGATG 391
Db      194  AAAACAGCGTGTATTTATTCACAAAGATTTCTTACCAACATCAAAACCGGCTGTGT 253
Oy      392  GGAATATCAAGAAAGTGGAGACAAAGTGAATTTTGGCTTTTAATGCTGATCCAA 451
Db      254  TTAATTTCAAGCCTAAGCAATTTAATAATGATTAAGACAGAGCGGAATTTCTATCTT 313
Oy      452  GATATTCATGCTTGGTGATCAAGAACTGAAAAGGAGACCTTCTGTACATGCA 511
Db      314  TAGATACAGATTTAGGCATCTTGTCTC---ATCTGAATCTGATCCTGAAATATCAT 370
Oy      512  GTTGGTGCTAGCCGACGAGAAATGACTGTGATCAGAGAAAGGTGATGATG 571
Db      371  GTATGTTGCTAGACATGCTGAGAGAGTGGCATGCTTTCAGAAAGAAATGTTTACATG 430
Oy      572  TGTATTTAAAGCAGCAGCTTTAGTGAAGAGCAGCTGAGAT---CTACCGATGATTTAG 628
Db      431  TTGTGCTCAAGCAGCAGCAGCTTGTATTAAGCAACCAAGTGGGTTACTGCTATCTTG 490
Oy      629  AAGTGAACCGGATATTTGTTTCCAGGGAATTTGGCTTGTGATGAGACATATGATC 688
Db      491  ATGGAACCCAGATATTTGCTTACCCGGAATTTAAGTGTGCTCAGTGAAGCTTATGATC 550
Oy      689  GTTGTGCGAAGTATGTCAGAGATATCCAAAGACATTTTCTTGAAGCAACCAAGTATGA 748
Db      551  GTTGTGAGAAATTTGGCCGAGATGCTAAGACATTTTACTTGGGAACTTTGCTGATGA 610
Oy      749  CCCAGAGAGAAAGAGACTATCGGCAATATATGTGTGTGTCAGAGAAACGATAGC 808
Db      611  CTTCTGAAGCCGAGAGGCTATATAGGCTATATATGTGTGTGATGAGAGACATGAGC 670
Oy      809  TTGTTGATGCGCTTAATGATCATCCCAATTAATCTCGCAAGCTTTAGATGAGTGAGACA 868
Db      671  TCGTTGATGGGCTTAATGCTTCAACATACTCCAAACAGCTTTAGACAGGTGGAGTCCA 730
Oy      869  GCGTGAAGATATTTCAAGTGGGGGCGCATTTGATATGCTTGAAGCTTATCCGATA 928
Db      731  GGTGGAAGACCTTTTCGGGGGTGTCATTGTAATGATGCTTGAATTAACAATA 790
Oy      929  CTGTCTCAAGATTTCCGTTGATATTCAGCCATTCAGAGATATGATGAAGAAATGCGTA 988
Db      791  CAGTACCAAAATTTCCGTGACATTCAGCCATTCAGAGATATGATGAAGAAATGAGGA 850
Oy      989  TGAATCTGTGAATTCAGATCAAAACCTTGCATGATGATCTATTTGTTAGTATG 1048
Db      851  TGAATCTGTGAATTCAGATCAAAACCTTGCATGATGATCTATTTGTTAGTATG 910
Oy      1049  TTGCTGTATCTGTAGATATGATGATGCTTCAAGTATGAGGATATGCACTGAATCAAGG 1108
Db      911  TTGCTGTAGACCGTATGAGCTATATGATGCTTCAAGTATGAGGATATGCACTGAATCAAGG 970
Oy      1109  CAAACAAGAGAGATGATATATCTCTTGTGCTTTAGGCTTGAACATCACTAACA 1168
Db      971  CAAACAAGAGAGGCTTCAATGATGACATTTGGCATAGGATTTGCTAATCAGTCACTA 1030
Oy      1169  ATATACTCAGATGTAGAGAAAGATGCCAGAAAGAGAGAGATATCTTGCTCAAGATG 1228
Db      1031  ACATACTCAGAGATGTTGAGAGAGATGCCAGAAAGAGAGAGGTTTATCTACACAAATG 1090
Oy      1229  AATTAGACAGGAGGAGCTTCCGACAGAAACATATTTGCTGGAAGAGTACATGATAGT 1288
Db      1091  AGTTGGACAGGAGGAGCTTCCAGATATGATATGACATATTTGCTGAGAGAGTACATTAAT 1150
Oy      1289  GGAAGAACTTATGAAGAAACAAATTCAGAGGAGGAGAAATTTGATGATGATGACAGGA 1348
Db      1151  GGAAGAACTTATGAAGAAACAAATTTAAGAGGAGGAGAAATTTGATGATGATGACAGGA 1210
Oy      1349  AAGGTGTCAGAACTGCACTGCTATGATAGTGGCTGTGTTAAACAGCGCTGCTGTGT 1408

```

Db 1211 ACGGTGTACCGAGCTGAGTGAAGCTAGTGCAGTGGCGGTGAGGCTTATGCTGTGT 1270

Qy 1409 ATGCGAAGATATTGACGAGATTGAAGCCAGATACAACTTACAGAGAGGGCTT 1468

Db 1271 ACCGGGAATTAATGATGATGATGAGCCAAATGATTAACAACCTTCAAGAGAGGCTT 1330

Qy 1469 ATGTAGCAAGCCAAAGAACTTCTGACCTTGGCCATGCTTAATGCAAAATCTTGTGC 1528

Db 1331 ATGTAGTAAGACCAAGAGATAGTGCATACCAATGCAATGCAATGCAAAATCCCTTAC 1390

Qy 1529 CCCCTAATGAA 1540

Db 1391 GCCCGTCAAGAA 1402

RESULT 13

AB037975 1689 bp mRNA linear PLN 25-NOV-2000

LOCUS AB037975

DEFINITION Citrus unshiu mRNA for phytoene synthase, complete cds.

ACCESSION AB037975.1 GI:11344506

VERSION phytoene synthase.

KEYWORDS Citrus unshiu

SOURCE Citrus unshiu

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Sukuariophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Sapindales; Rutaceae; Citrus.

REFERENCE 1 (sites)

AUTHORS Ikoma,Y., Komatsu,A., Kita,M., Ogawa,K., Omura,M., Yano,M. and Moriguchi,T.

TITLE Expression of a phytoene synthase gene and characteristic carotenoid accumulation during citrus fruit development

JOURNAL Physiol. Plantarum 111, 232-238 (2001)

AUTHORS 2 (bases 1 to 1689)

AUTHORS Moriguchi,T.

REFERENCE Direct Submission

JOURNAL Submitted (04-FEB-2000) Takaya Moriguchi, National Institute of Fruit Tree Science, Department of Research Planning and Coordination; 2-1 Fujimoto, Tsukuba, Ibaraki 305-8605, Japan (E-mail: takaya@fruit.affrc.go.jp, Tel:81-298-38-6416, Fax:81-298-38-6437)

FEATURES

source location/Qualifiers

1..1689

/organism="Citrus unshiu"

/mol\_type="mRNA"

/cultiivar="satsuma mandarin"

/db\_xref="taxon:55188"

/issue\_type="pulp"

/dev\_stage="mature stage"

179..1489

/codon\_start=1

/product="phytoene synthase"

/protein\_id="BAB18514.1"

/db\_xref="GI:11344507"

/translation="MSVTLLVWSPNSQSNCFEVDVSVRENRLLFYSSRFYLOHQR  
TAVPNSRPKQFNNSKORNSYPLDLDLPCSSGSDIDPEISLSEAYDSCGVCAYATPY  
WYVNVVVKQALVKNQPSGVTBDLDVNPDIAPGTLISLSEAYDSCGVCAYATPY  
LGTLLMSTRRAIWAIVYKCRTEDELVDPNASHITTPALRMESRLDELFRGPFD  
KIDALSDTVKFPVDIOPFRMIEGRMDLRSRKXNDELILYCYVAGVGMV  
PVMGIAPDSQATTESVYNAALAGIANOLTLIRVDGEDARGRVYLPDDELQAQGLS  
DDDIPIAGEVTIKRNFMKNOIKRAMFFPMANGVTEISEASRMFVMSALLYRQILD  
EIEANDYNNFTRAYVSKAKITAPLPIAVAKSLRPSRLYTSKA"

CDS

BASE COUNT 512 a 302 c 388 g 487 t

ORIGIN

Query Match 40.3%; Score 696.8; DB 8; Length 1689;

Best Local Similarity 71.2%; Pred. No. 2.1e-159;

Matches 949; Conservative 0; Mismatches 377; Indels 6; Gaps 2;

Qy 212 AATGTTTAGACACCGAGGTTCTGTTTCAATGACATGCTGTTGCTTTGTGTGG 271

Db 141 AATATTTCTTTTATCCAAATAATATTTTACAGCATGCTGTATGATGCTGTGG 200

Qy 272 TTGTTTCTCCACCTTCCGAGGTCTCGAATGGACAGATTTGTGATTCAGTCCGAGAG 331

Db 201 TTGTATCACTTAACCTTACCAATTTGCTTGGGTTGGTTCATTCAGTTCGAGAG 260

Qy 332 GAAACCGGCTCTTTGATATCAGGTTCCTAGCTCGAGATAGGAATTTGATGTGAATG 391

Db 261 AAAACAGGCTGTTTATTCATCAAGATTTCTTACCAATCAACCCGAGCTGTGT 320

Qy 392 GGAGAAATCAAGAAAGGTGGAGACAAAGTGAATTTTGGCTTTAATTCGTATCA 451

Db 321 TTAATTTAGACTTGAAGATTTAATTAATTAATTAATTAATTAATTAATTAATTA 380

Qy 452 GATATTTAGCTTGGGTGATCAAGAACTGAAAGGAAGCACTTCTCTGATACGTCA 511

Db 381 TAGATACAGATTTGAGGATCTTGTCTC---ATCGAATGACCTTGTGTAATATCAT 437

Qy 512 GTTGGTGCTTACCCAGCTGGAAGAAATGACTGTCTATCAGAGAAAAAGTGTATGAT 571

Db 438 GTATGTTGCTTACGCTCTGAGAAAGTGGCCATGTCTTCAAGAAATGTTTCAATG 497

Qy 572 TGGTATTAAGCAGGACCTTATGTAAGAGGACGTGAGAT---CTACCATGATTTAG 628

Db 498 TTGCTCTAAGCAGGACGCTTGTGTAATTAATTAATTAATTAATTAATTAATTAAT 557

Qy 629 AAGTGAACCCGATATTTGTTTCCAGGGAATTTGGCTTGTGATGAGCATATGATC 688

Db 558 ATGTGAACCCGATATTTGCTTTACCCGGAATTTAAGTCTGCTGATGAGCTTATGATC 617

Qy 689 GTTGGCGAAGATATGTCAGAGATATGCAAGAACTTTTACTTAAAGCAAGCTAATGA 748

Db 618 GTTGGGAGAAAGTTTCCCGCGGTATGCTAAGACATTTTAACTTGGGAATTTGTGATGA 677

Qy 749 CCCGAGAGAAAGAAAGACTATCTGGGCAATTAATGTTGGTGGAGAGAAAGATGAGC 808

Db 678 CCTGAAAGGCGAAGGCTATATGTTGGCTAATTAATGTTGGTGGAGAGACAGATGAGC 737

Qy 809 TTGTTGATGGCCCTATGATCATCCCATATCTCCCAAGCTTTAGATAGTGGAGACA 868

Db 738 TCGTTGATGGGCTATATGCTTACACATATCTCCCAAGCTTTAGACAGTGGAGTCCA 797

Qy 869 GGTGGAAGATATTTTCAAGTGGGCGGCAATTTGATATGCTGATGCTGTTATCCGATA 928

Db 798 GGTGGAAGACCTTTCCGGGGTGTGCTATTTGATATGCTGATGCTGATATCAGATA 857

Qy 929 CTGTCTCCAGATTTCTGTTGATATTCAGCCATTTAGAGATATGTTAGAGAAATGCGTA 988

Db 858 CAGTACCAAAATTTCTGTGCAATTCAGCAATTCAGAGATATGATTAAGAGAAATGAGGA 917

Qy 989 TGAATCTTGGAAATCCAGATPACAAACCTTGCATGAGCTATATGCTATTTGTTACTATG 1048

Db 918 TGGAATCTTAGAGATCAAGATPACAAACCTTGCATGAGCTATATGCTATTTGTTACTATG 977

Qy 1049 TTGCTGTAGTGTAGATGATGATGATGCTTCCAGTTATGGGTAATTCACCTGATCAAAAG 1108

Db 978 TTGCTGGAGACGTAGGCTTAATGATGTTTCCAGTTATGGGCAATACCTGATCCTACAGG 1037

Qy 1109 CAACAAAGAGATATATATGCTGCTTGGCTTTAGGCGCTTCAAAATCACTTACCA 1168

Db 1038 CAACAAAGAGAGCTCTTACATGACATGTCAGTCTGAGGATTTCTTACAGCTACCTCA 1097

Qy 1169 ATATCTCAGAGATATGAGAAAGATGTCAGAAAGAGAGATATGCTGCTCAAGATG 1228

Db 1098 ACTATCTCAGAGATTTGAGAGATGTCAGAAAGAGAGATTTATCTTACCAAGATG 1157

Qy 1229 AATTAGCAGGACGAGGCTTCTCCAGAAAGATATTTGCTGGAAGATGACTGATAGT 1288

Db 1158 AGTTGACAGGACGAGGCTTCTCAGATATGATGATATTTGCTGGAAGATGACTGATAGT 1217

Qy 1289 GGAGAACTTTATGAGAAACAAATTCAGAGGCGAGAGAAATTTCTTATGATGATGAGGA 1348

Db 1218 GGAGAACTTTATGAGAAACAAATTTAGAGGCGAGAGATTTCTTATATGCTGAGGA 1277

Qy 1349 AAGTGTACAGAACTGACTCTGTATGATGAGCTGTGTAAACGCGCTGCTGTGT 1408

[illegible]

OY	300	TGGGACAGAGATGTGGATTCAGTCGCCAAGAAAGAAACCGCGTCTTTGTATCATCCAGTT	359
Db	301	TTGGCTTGGGGTTCGTCAATTCAGTTCAAGAGAAACAGCGCTGTTTATTCATCAAGATT	360
OY	360	CTTAGCTCGAGATTTGGAATTTGATGTGATGGAGAAATCAAGAAAGGTGGAGACAAG	419
Db	361	TCCTTACCAACATAAACCCGAGCTGCTGTGTTTAATTCAGACTAAGCATTTAATA	420
OY	420	GTGGAAATTTTGGCTCTTTAAATTCGTGATCCAGATATTCATGCTTGGGTGATCAAGAAC	479
Db	421	TAGTATATAGAGAGACGGAATTCCTTATCCTTTAGATACAGATTTGAGCATCTTCTCTC	480
OY	480	TGAAAGGAAAGCACTTCTCTGTACAGTCCAGTTGGTGGCTTAGCCCACTGGAGAAAT	539
Db	481	---ATCTGAATTCACATCGCTCGTGAATAATCATGTATGTTGTCTAGCACTCTGTGAGAGT	537
OY	540	GACTGTGCATCAGAGAAAAAGGATATGATGTGATATTAAGACAGGACAGCTTATAGGAA	599
Db	538	GGCATGTCTTTCAGAAAGAAATGTTTCAATGTGTGTCTCAAGCAGGCAAGCTTTGTTAA	597
OY	600	GAGGCAAGCTGA--GATCTACCGATGATTTAGAAGTGAAGCCGGATATTTGTTTCAGG	656
Db	598	TAAACAAACCAAGTGGGTTACTCGTGAATCTTGATGTGAACCCAGATATTTGCTTACC	657
OY	657	GAATTTGGCGCTTTTGATGTGAAGATATGATCGTTGTGGGAAATATGTCAAGATATGC	716
Db	658	AACCTTAAGTCTGCTCAGTGAAGCTTATGATCGTTGTGGAAAGTTTGCCCGAGTATGC	717
OY	717	AAAGACATTTTACTTAGGAACCAAGCTAATAGACCCAGAGAGAAAGAGCTATCTGAGGC	776
Db	718	TAAACATTTTACTTGGAACTTTGCTGATGACCTCTGAAGAGGAAAGGCTATATGGC	777
OY	777	AATATATGTGTGTGACAGAAACGATGAGCTTGTGATGCGCCCTAATGATCCACAT	836
Db	778	TATATGTGTGTGTGAGAGACAGATGAGCTGTTGATGGCGCTAATGCTTACACAT	837
OY	837	AACCTCCGAACTTAAATATGATGTGGAGACAGGCTGGAAAGATTTTCAATGGGCGGCC	896
Db	838	AACCTCAAAGCTTTTAAACAGGTGGAGATCAGGTTTGAAGAACCCTTTTCCGGGGTCAAGC	897
OY	897	ATTGATATGTGTGAATGCTGCTTATACCGATACGTGTCCAGATTTCTGTGATATCA	956
Db	898	ATTGATATGTGTGTATCTGTGATATATGATACGTAACGTAACCAATTTCTGTGACATTTCA	957
OY	957	GCCATTCAGAGATATGATTAAGAGATGCGATGAGCTTGTGAAATCCAGATCAAAAC	1016
Db	958	GCCATTCAGAGATATGATTAAGAGAAATGAGAGATGAGACTTAGGAAGTCAAGATCAAAAA	1017
OY	1017	TTTCGATGAGCTATATCTATTTCTATGTGTCTGTGATCTGTAGAGATTGATGAGTGT	1076
Db	1018	CTTTGATGAATTAATACCTGTATGTTTATGTGTGTGGGACCCCTAAGGGCTAAAGAGTGT	1077
OY	1077	TCCAGTATAGGGATATGACCTGAAATCAAGGCAACAAAGAGATATATATATATGCTGC	1136
Db	1078	TCCAGTATAGGGATATGACCTGCTCAAGGCAACAAAGAGAGCTTATCAATATGACGC	1137
OY	1137	TTTGGCTTTAAGGGCTTCCAAATCAACTAACCAATATATCTCAGAGATGTAGAGAGATGC	1196
Db	1138	ATTGCACTAGAGAGATTCCTATATACGCTCACTAATATCTCAGAGATGTGGAAGAGATGC	1197
OY	1197	CAGAAAGAAAGATATACCTTGCTCAAGATGAATTAGCACAGGCAAGGCTCTCCAGCA	1256
Db	1198	CCCAAGAGAAAGGCTTTATCTACCACAAGATGAGTTGGCACAGGCAAGGCGCTTTCAGATGA	1257
OY	1257	AGACATATTTTCTGGAAGAGATCTGATTAAGTGAAGAACTTTATGAAAGAACAAATTC	1316
Db	1258	TGACATATTTTCTGGAAGAGATCAATTAATGAGAACTTCATGAAAGAACAAATTTA	1317
OY	1317	GAGGCGACAGAAATTTCTTATGATGTAGAGAAAGGTGTCAAGAACTGTGACTCTGTAG	1376
Db	1318	GAGGCGAAGATGTCTTTTATATGTGTGAGAAAGGTGTCAAGAGCTGTAGTGAAGCTAG	1377
OY	1377	TAGATGCGCTGTGTAAACAGCGCTGCTGTGTATGCAAGATATTGACGAGATTGAGC	1436

Dd	1378	TGAGTGGCCGGATGGCGTTCACTTGCTGTTCACCGGAATTACTGCATGATGATTAAGC	1437
Oy	1437	CAGACATTACAACAACCTTCACAGAAGGCGTTAATTGCAAGCCAAGAAGCTTTCAC	1496
Dd	1438	CAATGATTATACAACAACCTTCACAAAGAGCGTTGTGTGATAAGGCCAAGAAGATAGTCGC	1497
Oy	1497	CTTGCGCCATTTGGCTTATGCAAAAATCCTTGTCGCCCTTAATAGAA	1540
Dd	1498	ACTACCATTGCAATATGCAAAATCCCTCTTAGCCCCGTCAAGAA	1541
RESULT 15			
E15683			
LOCUS	E15683	2868 bp	DNA linear PAT 28-JUL-1999
DEFINITION	Gentiana lutea mRNA for Phytoene synthase 4, complete cds.		
ACCESSION	E15683.1	GI:5710366	
VERSION	JF 1998084966-A/4.		
KEYWORDS	unidentified		
SOURCE	unclassified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 2868)		
AUTHORS	Koiwa,H., Kobayashi,H. and Yamamura,S.		
TITLE	NEW PHYTOENE SYNTHASE GENE		
JOURNAL	Patent: JP 1998084966-A 4 07-APR-1998;		
COMMENT	IMATE PREP GOV		
	OS Gentiana lutea		
	PN JP 1998084966-A/4		
	PD 07-APR-1998		
	PF 17-SEP-1996 JP 1996245107		
	PI KOIWA HIROYUKI, KOBAYASHI HITOSHI, YAMAMURA SABURO PC		
	CI:CN15/09,C07H21/04,C12N9/00//C09B61/00,(C12N9/00,C12R1;91); CC		
	strandedness: Double;		
	CC Topology: Linear;		
	FH Key Location/Qualifiers		
	FH source 1..2868		
	FT /organism='Gentiana lutea'		
	FT /molecule='genomic DNA'		
	FT CDS 400..1689		
	/product='phytoene synthase 4'.		
FEATURES			
source	location/Qualifiers		
	1..2868		
	/organism="unidentified"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32644"		
BASE COUNT	887 a 499 c 648 g 833 t 1 others		
ORIGIN			
Query Match	40.1%; Score 692.4; DB 6; Length 2868;		
Best Local Similarity	73.6%; Pred.No. 2.4e-159;		
Matches 949; Conservative	0; Mismatches 311; Indels 30; Gaps 4;		
Oy	240 TTTCATGAGCATGTCTGTGTGCTTGTGTGGTGTGTTCCTGCCACTTCAGAGTTCGAA	299	
Dd	390 TTCTGTATACATGTCTATTTGTACCTATGGGTCTTTCGCCAGTTCGAAGTTTTAG	449	
Oy	300 TGGGACAGATTTGTTGATTCAGTCGCGAAGAAACCGCGCTTTGATATCATCAGATT	359	
Dd	450 TGGCATGTTTCTTGAGGCCAATTCAGAAAGTTTACCATTTTTCG-----	495	
Oy	360 CCTAGCTGAGATAGAAATTTGATGTGGAATGGAATCAAGAAAGGTGGAGACAAG	419	
Dd	496 -----GATTAAGTTTATGTACATGGAAGAGTTAAGAAAAATGACACCAAAG	545	
Oy	420 GTGCAATTTTGGCTTTTAATTGCTGATCCAAAGATATTCATGCTTGGGTGATCAAGAC	479	
Dd	546 GCCTGATCATCAGTTATGGGGTTGGAGATTTGATTCATTTGCTTGAAGAGATCTGAGTT	605	
Oy	480 TGAAG---AGGGAACACATTTCTCTGTCAGTCGCGATTTGGTGCTAGCCAGCTGGAGA	536	
Dd	606 AGAGACCCCAGGGAAGAGATTTATGGGTATCTCTCACAGATTTATAGCTACCCCGGACGAGGA	665	

OY	537	AAATACGTGTGCATCAGAGAAAAAGGTATGTGATGTGATATTAAGCAGCAGCTTAGT	596
Db	666	AATATCAATGACATCAAGCAGAAAGTTTATGATGTGCTTTTAAGCAGACGCTTTGAT	725
OY	597	GAAAGGCAAGCTGAGATCTACCGATGATTTAAGATGAAGCCGAGATATGTGTTCCAGG	656
Db	726	TAAATAGACAGTTGAGGCTCTAGAGAAAAATTTGAGAGGTGAACCCGACATTAATTTGCCAGG	785
OY	657	GAATTTGGCGCTGTGTGAGTGAAGCATATGATTCGTTGTGGCAACTATGTGACAGGTATGC	716
Db	786	AAAGCCGAACGCTGTGATGATGAAGTTATGATCGGTGTGCAAGATATGTGCTGAATATGC	845
OY	777	AATATATGTGTGTGCAGAGAGACGATGAGCTTTGTGATGTGACCTTAATGCAATCCACAT	836
Db	837	AACCTCCCAAGCTTTAATATAGTGGGAGACCAAGCTGTGAAGATATTTTCAGTGGGCGCGC	896
OY	966	AAATTCACACCGCTTAAATAGTGGGAAACAAAGATTGAAGATGTTTTCAAAGGGCAACC	1025
Db	897	ATTTGATATGCTGTATGCTGCTTTATCCGATCTAGTCTCCAGATTTTCTGTGTGATATYCA	956
OY	1026	TTTGTGATATGTGTATGCTGTCTTTATCTGATACCATTAACCAAGTATCTGTGTGACATCCA	1085
Db	957	GCCATTCAGAGATATGATTTGAAGGAATGCGTATGGACTTTGTGAAATCCAGATACAAAC	1016
OY	1086	GCCATTTAGAGATATGATATGAAGGAATGCGGATGATCTGAAAGAAATCGAGATCAAGAA	1145
Db	1017	TTTGATGAGCTATATCTATTTGTATCTATGTTGCTGGTACTGTAGATTTGATGAGAGT	1076
OY	1146	TTTGGATGAGCTGATCTTATCTGCTATATATGTGGCTGTGATCAAGTGTGCTTGAATGAGTGT	1205
Db	1077	TCCAGTTATGGGATATGSCAACCCTGAATCAAAAGCAACAAGAGTATATATATATGCTGC	1136
OY	1206	ACCAGTATATGGGCAATTCGACCTGAACTTAAAGGCAACAAGAAAGTGTATATATGCAAGC	1265
Db	1137	TTTGGCTTTAGGGCTTGCAAAATCAACTAACCAATATATCTCAGAGATGTAGAGAAATGC	1196
OY	1266	TTTATCTTTTGGGATGCGGAACCAAGCTGACTTCAATCTTAAAGGATGTGTGAGAAAGTGC	1325
Db	1197	CAGAAAGAAAGATATATCTGCTCAAGATGAAATTTAGCACAGCAGGCTCTCCGACGA	1256
OY	1326	AAGAAAGGAGAGATGTACTTAAGCTCAAGATGAATTTAGCAGACAGAGTTTATCAGATGA	1385
Db	1257	AGACATATTTGCTGGAAGAGTGACTGATTAAGTGGAGAACTTTATGAGAAACAAATYCA	1316
OY	1386	GGAATTTTGTGCTGGAAGAGTTTACAGACAAATGAGAGATTTTATGAGAAAGCAATCA	1445
Db	1317	GAGGGCGAGAAATTTCTTTATATGATGATGAGAGAAAGGTGTCAAGAACTGCACTGTGC-TA	1375
OY	1446	AAGGCTTATGAAATATCTATGATGATGAGAGAAAGGAGTCCCGCACTCAAGCTCCGGAA	1505
Db	1376	GTATATGGCGCTGTGTTAAC--AGGCTGCTGTTGTATGCAAGATATTTGACGAGATYGA	1433
OY	1506	GCAATTTGCCGTGTGTGGGCGAGCGTTTGCTTTTATATGAAAAATATTTGATGAGATATGA	1565
Db	1434	AGCCACGACTACAACACTTACACAGAGGGCTTATGTTAGCAAGCCAAAGAAAGCTTCT	1493
OY	1566	AGCAAAATGACTACAACAATYTCACAAAAAGGGCTTATGATATAAACAAAGCAAAAGATATY	1625
Db	1494	CACCTGGCCATGCTTATGCAAAATCTCT	1523
OY	1626	AGCTATGCTGTACATATGCTCAAGTCTCT	1655



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 00:00:49 ; Search time 480 Seconds  
(without alignments)  
9717.977 Million cell updates/sec

Title: US-09-847-081B-1

Perfect score: 1728

Sequence: 1 agaaccaccagaagacacac.....tcaccaaacctcaagtga 1728

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq\_19Jun03:\*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1728	100.0	1728	24	AA166366
2	1405.6	81.3	1573	25	ABX13568
3	868	50.2	1826	19	AAV03880
4	858.6	49.7	1795	19	AAV03878
5	857.8	49.6	1712	24	AA166367
6	853.4	49.4	1814	19	AAV03881
7	847.6	49.1	1591	17	AAQ9323
8	840.4	48.6	1646	12	AAQ12495

9	811.2	46.9	1316	19	AAV03879
10	806.4	46.7	1239	21	AAZ39482
11	806.4	46.7	1239	24	AAZ40272
12	692.4	40.1	2868	19	AAV16951
13	678.4	39.3	1921	19	AAV16948
14	670.4	38.8	2085	19	AAV16948
15	653.4	37.8	1932	19	AAV16950
16	578.6	33.5	1703	21	AAV16950
17	578.4	33.5	1269	24	ABZ12924
18	578.4	33.5	1566	21	AAV16950
19	548.2	31.7	1397	21	AAZ29145
20	512.2	29.6	1304	24	ABA97361
21	456.8	26.4	992	21	AAZ29144
22	424	24.5	1448	21	AAZ29139
23	415.2	24.0	1239	19	AAV17247
24	401	22.2	1021	21	AAZ29146
25	388.8	22.5	1060	21	AAZ29143
26	290.6	16.8	2585	23	ABL41600
27	251.4	14.5	888	21	AAZ29140
28	223.8	13.0	3485	13	AAQ29121
29	177.4	10.3	684	22	AAH44248
30	172	10.0	749	19	AAV03882
31	170.2	9.8	476	21	AAZ29142
32	145.6	8.4	1386	13	AAQ29122
33	121.8	7.0	766	21	AAZ29141
34	82.6	4.8	879	24	ABO82333
35	59.2	3.4	5609	21	AAH41459
36	59.2	3.4	349980	21	AAF21609
37	59.2	3.4	143768	21	AAH41450
38	55.6	3.2	870	25	ABZ40269
39	48	2.8	4590	22	AAH4065
40	43.2	2.5	14041	22	AAH48024
41	43.2	2.5	440375	22	AA199683
42	43.2	2.5	4411529	22	AA199682
43	42.2	2.4	948	19	AAV73180
44	41.8	2.4	332	19	AAV73190
45	41.6	2.4	102	21	AAZ59432

## ALIGNMENTS

RESULT 1	AA166366	strand: cDNA, 1728 BP.
ID	AA166366	
AC	AA166366	
XX		
DT	29-JAN-2002	(first entry)
XX		
DE	Nicotiana tabacum phytoene synthase coding sequence #1.	
XX		
KW	Phytoene synthase; zeta carotene desaturase; herbicide; transgenic plant;	
KW	plant growth regulator; herbicide; tobacco; ss.	
XX		
OS	Nicotiana tabacum.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	244..1566
FT		/tag= a
FT		/product= "phytoene synthase"
XX		
PN	DE10022362-A1.	
XX		
PD	15-NOV-2001.	
XX		
PF	08-MAY-2000; 2000DE-1022362.	
XX		
PR	08-MAY-2000; 2000DE-1022362.	
XX		
PA	(FARB ) BAYER AG.	
XX		
PI	Busch M, Hain R;	

Phytoene synthase  
cDNA encoding a ph  
tomato phytoene sy  
Nucleic acid encod  
Nucleic acid encod  
Nucleic acid encod  
Nucleic acid encod  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Soybean phytoene s  
Nucleotide sequenc  
Soybean phytoene s  
Corn phytoene synt  
MTOM5 modified ph  
wheat phytoene syn  
Rice phytoene synt  
pay A encoding seq  
Corn phytoene synt  
Fragment of GTOM5  
Physcomitrella pat  
Phytoene synthase  
Rice phytoene synt  
Fragment contg. pt  
Rice phytoene synt  
Human NOV8 encodin  
N. meningitidis pa  
Neisseria meningit  
N. meningitidis B  
N. gonorrhoeae nuc  
Yeast AOP9604-asso  
Internal control B  
Mycobacterium tube  
Mycobacterium tube  
C. utilis crtB DNA  
Sample Spirulina n

XX WPI, 2002-02736/04.  
 DR P-PSDB; AAM51841.  
 XX  
 PT New nucleic acid encoding tobacco zeta-carotene desaturase, useful for  
 screening compounds with herbicidal activity -  
 PS  
 XX Claim 14; Page 12-17; 44pp; German.  
 XX  
 CC The present invention provides the protein and coding sequences of  
 CC phytoene synthase and zeta-carotene desaturase from *Nicotiana tabacum*.  
 CC The sequences can be used to identify compounds capable of altering the  
 CC expression of these genes, which are therefore useful as plant growth  
 CC regulators and herbicides. They can also be used to produce transgenic  
 CC plants. The present sequence is the coding sequence of a tobacco  
 CC phytoene synthase.  
 CC  
 XX Sequence 1728 BP; 513 A; 290 C; 433 G; 492 T; 0 other;  
 SQ  
 Query Match 100.0%; Score 1728; DB 24; Length 1728;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAACCCAGAAAGAACCAAGCTTTGCTTGTGATGAGTCAATTCCTGCTT 60  
 DB 1 AGAAACCCAGAAAGAACCAAGCTTTGCTTGTGATGAGTCAATTCCTGCTT 60  
 QY 61 GTGTAGGCAAGTGGCTTCACTTCTTATATCCGATTTTATATCGTTAAATTAATG 120  
 DB 61 GTGTAGGCAAGTGGCTTCACTTCTTATATCCGATTTTATATCGTTAAATTAATG 120  
 QY 121 GATAGACTCTAGTGTATCTACAGTATGTTGTTTGTATGATTAATAGCTGAGTGA 180  
 DB 121 GATAGACTCTAGTGTATCTACAGTATGTTGTTTGTATGATTAATAGCTGAGTGA 180  
 QY 181 AGTAAACATTAAGAAAGCAAAACCTTGGAAATGTTTATAGACCAAGGTTTCTTGT 240  
 DB 181 AGTAAACATTAAGAAAGCAAAACCTTGGAAATGTTTATAGACCAAGGTTTCTTGT 240  
 QY 241 TTCAAGACATCTGCTTGTGCTTGTGTTGGTGTGTTTCCCACTCCGAGGCTCCGAT 300  
 DB 241 TTCAAGACATCTGCTTGTGCTTGTGTTGGTGTGTTTCCCACTCCGAGGCTCCGAT 300  
 QY 301 GGGACAGATTTGTTGATTCAGTCCGAGAAAGAACCCGCTTGTATCAATCCAGTTC 360  
 DB 301 GGGACAGATTTGTTGATTCAGTCCGAGAAAGAACCCGCTTGTATCAATCCAGTTC 360  
 QY 361 CTAGCTCGAGATGAGAAATTTGATGGAATGGGAATCAAGAAAGTGGGAGACAAAG 420  
 DB 361 CTAGCTCGAGATGAGAAATTTGATGGAATGGGAATCAAGAAAGTGGGAGACAAAG 420  
 QY 421 TGGAAATTTTGGCTTTTAAATGCTATCCAAATATTCATGCTTGGTGAATCAAGAACT 480  
 DB 421 TGGAAATTTTGGCTTTTAAATGCTATCCAAATATTCATGCTTGGTGAATCAAGAACT 480  
 QY 481 GAAAAGGAGAGACATTTCTCTGTACAGTCCAGTTGTGCTAGCCGAGCTGAGAAATG 540  
 DB 481 GAAAAGGAGAGACATTTCTCTGTACAGTCCAGTTGTGCTAGCCGAGCTGAGAAATG 540  
 QY 541 ACTGTGTCATGAGAAAGGTGTATGATGTATTAAGCAGGACGCTTTAGTGAAG 600  
 DB 541 ACTGTGTCATGAGAAAGGTGTATGATGTATTAAGCAGGACGCTTTAGTGAAG 600  
 QY 601 AGGACGCTGAGATCTACCGATGATTTAGAGTGAAGCCGAGATTTGTTCCAGGGAT 660  
 DB 601 AGGACGCTGAGATCTACCGATGATTTAGAGTGAAGCCGAGATTTGTTCCAGGGAT 660  
 QY 661 TTGGCTTGTGAGTGAAGCATATGATGTTGGCGAAGTATGTGCAAGATATGCAAG 720  
 DB 661 TTGGCTTGTGAGTGAAGCATATGATGTTGGCGAAGTATGTGCAAGATATGCAAG 720  
 QY 721 ACATTTTACTTAGAACCAAGCTAATGACCCAGAGAGAAAGAGCTATCTGGGCAATA 780  
 DB 721 ACATTTTACTTAGAACCAAGCTAATGACCCAGAGAGAAAGAGCTATCTGGGCAATA 780

DB 721 ACATTTTACTTAGAACCAAGCTAATGACCCAGAGAGAAAGAGCTATCTGGGCAATA 780  
 QY 781 TATGTGTGAGAGAGAGAGATGAGCTTTGATGATGAGCCCTAATGCAATCCATAC 840  
 DB 781 TATGTGTGAGAGAGAGAGATGAGCTTTGATGATGAGCCCTAATGCAATCCATAC 840  
 QY 841 CCGCAAGCTTTAGATGAGTGGAGACCAAGCTGGAGATATTTTCACTGGGCGGCAAT 900  
 DB 841 CCGCAAGCTTTAGATGAGTGGAGACCAAGCTGGAGATATTTTCACTGGGCGGCAAT 900  
 QY 901 GATATGCTTGAATGCTCTTTATATCCGATCTGTCTCCAGATTTTCTGTTGATAT 960  
 DB 901 GATATGCTTGAATGCTCTTTATATCCGATCTGTCTCCAGATTTTCTGTTGATAT 960  
 QY 961 TTGAGATATGATTAAGAGATCCGATGAGATCTTGGAAATCCAGATCAAAATTTTC 1020  
 DB 961 TTGAGATATGATTAAGAGATCCGATGAGATCTTGGAAATCCAGATCAAAATTTTC 1020  
 QY 1021 GATGAGCTATATCTTATTTGATATGATGTTGCTGTGATCTGATGATGATGATG 1080  
 DB 1021 GATGAGCTATATCTTATTTGATATGATGTTGCTGTGATCTGATGATGATGATG 1080  
 QY 1081 GTTATGGGATATGCACTGATCAAGGCAACACAGAGTGTATATATGCTGCTTTG 1140  
 DB 1081 GTTATGGGATATGCACTGATCAAGGCAACACAGAGTGTATATATGCTGCTTTG 1140  
 QY 1141 GCTTATGGGCTTGGCAATCAACTAACATATATCTGAGATGTAGAGAGATGCCA 1200  
 DB 1141 GCTTATGGGCTTGGCAATCAACTAACATATATCTGAGATGTAGAGAGATGCCA 1200  
 QY 1201 AGAGAAAGATATCTTCTCTCAAGATGATTAAGCAGGACGAGCTCTCCAGAAAG 1260  
 DB 1201 AGAGAAAGATATCTTCTCTCAAGATGATTAAGCAGGACGAGCTCTCCAGAAAG 1260  
 QY 1261 ATATTTCTGGAAGAGTGAATGATGATGAGAGAACTTTATGAGAAACAAATTCAG 1320  
 DB 1261 ATATTTCTGGAAGAGTGAATGATGATGAGAGAACTTTATGAGAAACAAATTCAG 1320  
 QY 1321 GCGAGAAATCTTGTATGATGATGAGAGAAAGTGTCAAGAACTGAGCTGTAGTGA 1380  
 DB 1321 GCGAGAAATCTTGTATGATGATGAGAGAAAGTGTCAAGAACTGAGCTGTAGTGA 1380  
 QY 1381 TGGCTGTGTTAACAGGCTGCTGTGATTCGAAATATTTGACGAGATTTGAAGCC 1440  
 DB 1381 TGGCTGTGTTAACAGGCTGCTGTGATTCGAAATATTTGACGAGATTTGAAGCC 1440  
 QY 1441 GACTACAACTTCAAGAGGCTTATGTTAGCAAGCCAAAGAGCTTCTCACTTG 1500  
 DB 1441 GACTACAACTTCAAGAGGCTTATGTTAGCAAGCCAAAGAGCTTCTCACTTG 1500  
 QY 1501 CCGATGCTTATGCAAAATCTTGTGCGCCCTAATGAACTTCTCCACTGAGAAAG 1560  
 DB 1501 CCGATGCTTATGCAAAATCTTGTGCGCCCTAATGAACTTCTCTCCACTGAGAAAG 1560  
 QY 1561 ACATGAATGAAGTATGATGATGATGATTTATACCTAAAGAACTCAGATCTTGT 1620  
 DB 1561 ACATGAATGAAGTATGATGATGATGATTTATACCTAAAGAACTCAGATCTTGT 1620  
 QY 1621 AATGAGATATCTTTTGTAAATGTATGATCAATCAAAAGTAAATGTAATTAAT 1680  
 DB 1621 AATGAGATATCTTTTGTAAATGTATGATCAATCAAAAGTAAATGTAATTAAT 1680  
 QY 1681 ATCTCTGTGAGAAATATTTTCTCACTCACTAATCAACCTCTCAAGTGA 1728  
 DB 1681 ATCTCTGTGAGAAATATTTTCTCACTCACTAATCAACCTCTCAAGTGA 1728

RESULT 2  
 ABX13568  
 ID ABX13568 standard; DNA, 1573 BP.  
 XX  
 AC ABX13568;  
 XX







QY 1358 CAGAACTGACTGCTGCTAGTACAGTCCCTGCTGTTAAACAGCCGCTGCTGTTGATGCAAGA 1417  
 Db 1397 CACAACTGAGCTCAGCTAGTATGAGTCCCTGATAGGGCATCTTTGCTGTTAGACCCGACGA 1456  
 QY 1418 TATTGGACGAGATTGAAGCCCAACGACTACAACTTACAAAGAGAGGCTTATGTTACCA 1477  
 Db 1457 TACTGAGAGAGTTGAGACCAATGACTACAACTTACAAAGAGAGCTTATGTTAGACA 1516  
 QY 1478 AGCCAAAGAGCTTCTACCTTCCCTGCTTATGCAAAATCTCTGTCGCCCTTAATA 1537  
 Db 1517 AACCAAAAGAGTATTTCTTACTATTTGCTTATGCAAAATCTCTGTCGCCCTTAATA 1576  
 QY 1558 GAACCT-TCCTCTCCACTAGCAAAAGACATGATGAAGTGAAGTGAATGATTTATAC 1596  
 Db 1577 GAACCTCTGTCACCTCTAGCTAGGACATGACATGATTTAATTAAGCAAGAAACA 1636  
 QY 1597 ACTAAGAAAGCTAGCTCTGTTAATGATATCTTT 1634  
 Db 1637 TATATTATTATTATTATTACTGTTAAAGAAAGAAATTT 1674

RESULT 4

AAV03878 standard; cDNA, 1795 BP.

AAV03878;

29-APR-1998 (first entry)

Phytoene synthase coding sequence from N. benthamiana.

Phytoene synthase; transgenic plant; enhanced carotenoid synthesis;  
 ultra violet absorber; food colour; ss.

Nicotiana benthamiana.

Location/Qualifiers

361..1594

/\*tag= a

US5705624-A.

06-JAN-1998.

27-DEC-1995; 95US-0579667.

27-DEC-1995; 95US-0579667.

27-DEC-1995; 95US-0579667.

27-DEC-1995; 95US-0579667.

27-DEC-1995; 95US-0579667.

27-DEC-1995; 95US-0579667.

27-DEC-1995; 95US-0579667.

27-DEC-1995; 95US-0579667.

27-DEC-1995; 95US-0579667.

27-DEC-1995; 95US-0579667.

27-DEC-1995; 95US-0579667.

27-DEC-1995; 95US-0579667.

27-DEC-1995; 95US-0579667.

27-DEC-1995; 95US-0579667.

Sequence 1795 BP; 577 A; 271 C; 434 G; 513 T; 0 other;  
 Query Match 49.7%; Score 858.6; DB 19; Length 1795;  
 Best Local Similarity 76.4%; Pred. No. 8e-235;  
 Matches 1230; Conservative 0; Mismatches 274; Indels 107; Gaps 10;

QY 141 TACAAGTATGTTGTTTGTATTAATAGCTGAGGTGAGAAAGTAAACA-----TAAAG 193  
 Db 248 TAAATTTATTTATTTTATTTATTAATTAAGAGAGAGAGAGAGAAACGAAATCGAAAG 307  
 QY 194 GAAAGACAAAACCTTGGGAATGTTTATGACACCGAGGTTCTTGTTCATGACATGT 253  
 Db 308 TAAACAAAACCTTGGGAATGTTTATGACACCGAGGTTCTTGTTCATGACATGT 365  
 QY 254 CTGTGCTGTTGTTGAGGTTGTTCTCCCACTTCGAGAGTCTCGAATGAGACAGATGTT 313  
 Db 366 CTGTGCTGTTGTTGAGGTTGTTCTCCCACTTCGAGAGTCTCGAATGAGACAGATGTT 422  
 QY 314 TGAATTCAGTCCGAGAAAGAAACCGCGTCTTGTATCATCCAGGTTCTAGCTGAGATA 373  
 Db 423 TGAATTCAGTCCGAGAGAAACCGCGGTTTGTATGTTGTGAGGT-----ATA 470  
 QY 374 GGAATTCAGTCCGAGAGAAACCGCGTCTTGTATCATCCAGGTTCTAGCTGAGATA 433  
 Db 471 GAAATTTAGTGTGCAATGAGAGAAACAGAGAGTGGGAAACAAAGGTGAAATTTGG-- 528  
 QY 434 CTTTAATTCGATCCAGATATTCATGCTTGGGTGATCAAGAACTGAAAGGAGACA 493  
 Db 529 ----- 528  
 QY 494 CTTTCTGTAACAGTCCAGTGTGTTGTTGCTAGCCAGCTGAGAAAT---GACTGTGAT 550  
 Db 529 -----TTCTGTAAGTCTGATGTTGTTGCTAGCCAGCTGAGAAATGCGACATGAT 584  
 QY 551 CAGAGAAAAGGTTGATGATGTTGATTAAGACAGGCTTTTATGTAAGGCGACTGA 610  
 Db 585 CAGAACAGATGTTATGATGATGTTTAAACAGGCTTTTATGTAAGGCGACTGA 644  
 QY 611 GATCTACGATGATTTGAAGTGAAGCGGATATTTGTTGTTCCAGGAAATTTGGGCTTGT 670  
 Db 645 GATCTACGATGATTTGAAGTGAAGCGGATATTTGTTGTTCCAGGAAATTTGAGCTTGT 704  
 QY 671 TGAAGTGAAGTGAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 730  
 Db 705 TGAAGTGAAGTGAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 764  
 QY 731 TAGAACCAGCTATATGACCCAGAGAGAGAGAGAGAGTATCTGGGCAATATATGTGT 790  
 Db 765 TAGAACCAGCTATATGACCCAGAGAGAGAGAGAGAGTATCTGGGCAATATATGTGT 824  
 QY 791 GCAGAGAGAGAGAGAGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 850  
 Db 825 GCAGAGAGAGAGAGAGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 884  
 QY 851 TAGATAGGTGGAGACAGCTGGAAGATATTTTCAAGTGGCGGCAATTTGATATGCTTG 910  
 Db 885 TAGATAGGTGGAGACAGCTGGAAGATATTTTCAAGTGGCGGCAATTTGATATGCTTG 944  
 QY 911 ATGCTGCTTATCCGATATGTTCTCCAGATTTCTGTTGATATTTCAAGCATTCAGATA 970  
 Db 945 ATGCTGCTTATCCGATATGTTCTCCAGATTTCTGTTGATATTTCAAGCATTCAGATA 1004  
 QY 971 TGATTTGAAGAAATCGATATGAGCTTTGGAATTCAGATTAACAACTTGCATAGCATAT 1030  
 Db 1005 TGATTTGAAGAAATCGATATGAGCTTTGGAATTCAGATTAACAACTTGCATAGCATAT 1064  
 QY 1031 ATCTCTATTTGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1090  
 Db 1065 ACCTATATTTGTTATTTAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1124  
 QY 1091 TTGACCTGAATCAAGGCAACAAAGAGAGTATATATGCTGCTTTGGCTTTAGGGC 1150  
 Db 1125 TTGACCTGAATCAAGGCAACAAAGAGAGTATATATGAGCTTTGGCTTTGGGTA 1184

Claim 1; Column 15-20; 25pp; English.

DNA encoding tobacco phytoene synthase polypeptides - useful for  
 producing recombinant polypeptides or transgenic plants

This sequence encodes the phytoene synthetase from Nicotiana benthamiana.

The phytoene synthetase coding sequence represents a cDNA of the  
 invention. The isolated nucleic acid molecules are used for producing  
 recombinant polypeptides or transgenic plants with enhanced ability to  
 synthesise carotenoids. Phytoene has been used as a ultra violet absorber  
 and other carotenoids have been used as food colours, animal feeds and in  
 the pharmaceutical and cosmetics industries.



```

Db      854 CTTAGATAGGTGGGAACCGGCTTGAAAGTGTTCACGGGGGACCAATTGATAGCT 913
Qy      909 TGAATGCTTTATCCGATATCTGTCCAGATTTCCTGTGATATTCAGCCATTGAGAG 968
Db      914 CGATGCTGCTTTGTCCGATATCTGTTCAGAGTTTCCAGTTGATATTCAGCCGTTGAGAG 973
Qy      969 TATGATTTGAAGAAATCGTATGAGCTTGTGAAATCCAGATACAAAATTTCGATAGCT 1028
Db      974 TATGATTTGAAGAAATCGTATGAGCTTGTGAAATCCAGATATTAAGAACTTGTAGAGCT 1033
Qy      1029 ATATCTCTATGTTTACTATGTTGCTGTACTGTAGAGATTGATAGAGTTTCCAGTTATGG 1088
Db      1034 TTAACCTCTATGTTTATGAGCTGTGTGAGCTGTGTGAGTGTGATGATGTTTCAATTTAGG 1093
Qy      1089 TATTCACCTGATCAAGGCAACAAGAGAGAGTGTATATATGCTGTGCTTTGCTTTAGG 1148
Db      1094 TATTCACCTGATTCAGAGGCAACAAGAGAGCTGTATATATGACGCTTTGCTTTAGG 1153
Qy      1149 GCTTGCATATCACTAACCAATATCTCAGAGATGTAGAGAAAGATGCCAGAGAGAG 1208
Db      1154 AATCGCAATCACTAACCAATATCTCAGAGATGTGTGAGAGATGCCAGAGAGAGAG 1213
Qy      1209 AGATATCTGCTCAGAGATGATATAGCAGAGAGGCTCTCCGAGAGAGACATATTTGC 1268
Db      1214 AGCTATCTTACCTCAAGATGAATATACAGAGAGGCTCTTTCAGAGATGACATATTTGC 1273
Qy      1269 TGGAGAGTGAAGTATGATGAGAGAACTTTATGAAGAAACAATTCAGAGGGCGAGAA 1328
Db      1274 TGGAAAGTGAAGTATGATGAGAGAACTTTATGAAGAAACAATTCAGAGGGCGAGAA 1333
Qy      1329 ATTCTTGTAGTGAAGTGAAGAGAGTGTACAGAACTGGAATCTGTAGTATGAGGCTGT 1388
Db      1334 GTTCTTCATGAGGAGAGAGAGAGATTCACAACTGAGCTGATGAGAGATGAGCTGT 1393
Qy      1389 GTTAAAGAGGCTGCTGTGTGTATGCAAGATATGACAGATTTGAAGCCCAAGCTTACA 1448
Db      1394 ATGGGATCTTGTGCTGTGTGTACCGCAAAATACGAGAGATTTGAAGCCCAATGACTACA 1453
Qy      1449 CAATCTCAGAGAGAGGCTTATGTAGCAAGCAAGAAAGCTTCACCTGCGCCATTCG 1508
Db      1454 CAATCTCAGAGAGAGGCTTATGTAGCAAGCAAGAAAGCTTATTCCTTACTATTCG 1513
Qy      1509 TTAATGCAAAATCTTGTGCCCCCTATATGAACT-TCTCTCCATGAGCAAGACATGAA 1567
Db      1514 TTAATGCAAAATCTTGTGCCCCCTATATGAACTCTTGTACCTCTAGCTAAGCATAGA 1573
Qy      1568 TGAAGTAGTGAAGTCAATGATATTAACACTA---AAGAACTCAGTACTTGTAAAT 1623
Db      1574 CATCAGATTTAAATTAAGCAAGAAAGCATATCTGTAAAGAAAGAAATTTCTAAAG 1633
Qy      1624 GAGATATCTTT-----TGCTAAAGTGTATCATGAAAGTGAATGT- AATTCATAA 1675
Db      1634 TAGATATGTGTATGATGAGCCCTGTATATCATCAAAAGTAGTAGTAAATTCATAA 1693
Qy      1676 TGCAATCTCTTG 1688
Db      1694 TAACAATCTCTAG 1706

RESULT 6
AAV03881
ID      AAV03881 standard; cDNA; 1814 BP.
XX
XX      AAV03881;
XX
XX      29-APR-1998 (first entry)
XX
XX      Phytoene synthase coding sequence from N. tabacum.
DE
XX      Phytoene synthase; transgenic plant; enhanced carotenoid synthesis;
KW      ultra violet absorber; food colour; ss.
XX
XX      Nicotiana tabacum.
OS

```

```

XX      XX      Key      Location/Qualifiers
FH      CDS      363..1595
FT      CDS      /tag= a
XX      XX      US5705624-A.
XX      XX      06-JAN-1998.
XX      XX      27-DEC-1995; 95US-0579667.
XX      XX      27-DEC-1995; 95US-0579667.
XX      XX      27-DEC-1995; 95US-0579667.
XX      XX      (DELL/) DELLA-CIOPPA G R.
XX      XX      (FITZ/) FITZMAURICE W P.
XX      XX      (GRILL/) GRILL L K.
XX      XX      (HELL/) HELLMANN G M.
XX      XX      (KUMA/) KUMAGAI M H.
XX      XX      Della-Cioppa GR, Fitzmaurice WP, Grill LK, Hellmann GW;
XX      XX      Kumagai MH;
XX      XX      WPI; 1998-086196/08.
XX      XX      P-PSDB; AAW41060.
XX      XX      DNA encoding tobacco phytoene synthase polypeptides - useful for
XX      XX      producing recombinant polypeptides or transgenic plants
XX      XX      Claim 1: Column 33-36; 25pp; English.
XX      XX      This sequence encodes the phytoene synthetase from Nicotiana tabacum.
XX      XX      The phytoene synthetase coding sequence represents a cDNA of the
XX      XX      invention. The isolated nucleic acid molecules are used for producing
XX      XX      recombinant polypeptides or transgenic plants with enhanced ability to
XX      XX      synthesise carotenoids. Phytoene has been used as a ultra violet absorber
XX      XX      and other carotenoids have been used as food colours, animal feeds and in
XX      XX      the pharmaceutical and cosmetics industries.
XX      XX      Sequence 1814 BP; 577 A; 281 C; 441 G; 515 T; 0 other;
XX      XX      Query Match      49.4%; Score 853.4; DB 19; Length 1814;
XX      XX      Best Local Similarity 76.5%; Pred. No. 2.5e-233;
XX      XX      Matches 1223; Conservative 0; Mismatches 266; Indels 110; Gaps 10;
Qy      141 TCAAGTATGTTTGTGATTAATAGCTGAGTGAGAGG-----TAACTTAAG 193
Db      250 TAAATTTATTAATTTTATTAATTAAGCAGAGAGAGAGAGAAACAGAAACAGAAAG 309
Qy      194 GAAAGCAAAACCTGGGAATGTTTGAACACCGAGGTTTCTGTTTCATGAGCATGT 253
Db      310 TAAAGCAAAACCTGGGAATGTTTGAACACCGAGGTTTCTGTTTCATGAGCATGT 253
Qy      254 CTGTTGCTTGTGTGAGGTTGTTTCTCCACTTCGAGAGCTCGAATGGAACAGATTGT 313
Db      367 CTGTTGCTTGTGTGAGGTTGTTTCTCCACTTCGAGAGCTCGAATGGAACAGATTGT 313
Qy      314 TGAATTCAGTCCGAGAGAGAAACCCGCTTTGTATCATCCAGGTTTCTAGCTCAGATA 373
Db      424 TGAATTCAGTCCGAGAGAGAAACCCGCTTTGTATTCATCCAGGTTTCTAGCTCAGATA 373
Qy      374 GGAATTTGATGATGAGAGAGATCAAGAAAGTGGAGACAAAGAGTGAATTTGGCT 433
Db      472 GGAATTTGATGATGAGAGAGATCAAGAAAGTGGAGAGTGAATTTGGCT 433
Qy      434 CTTTAATGCTGATCCAGATATTCATGCTTGGGTGATCAAGAACTGAAAGGAGAGCA 493
Db      530 ----- 529
Qy      494 CTTTCTGTGATAGTCAAGTGTGTTGCTAGCCGAGCTGAGAAAT---GACTGTGAT 550
Db      530 ----TTCGTAAAGTCTGTGTATGTGAGCTACACCGCGAGAGAAATGCGAGATGACAT 585
Qy      551 CAGAGAAAAAGGTGATGATGTGTATTAAGCAGGACGCTTATGAAAGGAGCTGA 610

```

Db	586	CAGAACAGATGGCTTTATGATGTGGTTTTAAACAAGCAGCTTTAGTGAAGAGCAGTTGA	645
Oy	611	GATCTACCGATGATTTTAGAAGTGAAGCCGGATATTTGTTCCAGAGGAATTTGGCTTGT	670
Db	646	GATCTGATGATTTTAGAAGTGAAGCCGGAGATCCCTCTCCCGGGAATTTGACTTGT	705
Oy	671	TGAGTGAAGCATATGATGCTTTGGCGAAGTATGTGCAGAGTATGCCAAGAATTTTACT	730
Db	706	TGAGGAAGCAATATATAGTGTGTATGAGTATGTGCAGAGTATCCAAAGACATTTTACT	765
Oy	721	TAGSAACCAAGCTAATGACCCGAGAGAGAAAGAGCTATCTGGGCAATATATGTGTGT	790
Db	766	TAGGAACCATGTCTAATGACTCCAGAGAGAAAGAGGCTATTTGGGCAATATATGTGTGT	825
Oy	791	GCAGAGAACGGATGAGCTTGTGTATGGCCCTTAATGCAATCCCACTAACCTCCGAACTT	850
Db	826	GCAGAGAAACAGATGAACTTGTGTATGGCCCAAAAGCATCATGTATACACCCCAAGCT	885
Oy	851	TGATATGGTGGAGACACAGCTTCGAAGATATTTTCAATGGCGGCCCATTTGATATGCTTG	910
Db	886	TAGATAGGTGGGAGAACCGGCTTGAAGATGTTTTACGGGGGGACATTTGATATGCTCG	945
Oy	911	ATGCGCTTTATCCGATCTGTCTCCAGATTTCTCGTGAATTCAGTATTCAGCAATTCAGAGTA	970
Db	946	ATGCTGCTTTTGTCCCATCTGTTCCTCAGCTTCCAGTTGATATTCAGCCGTTTCAAGATA	1006
Oy	971	TGATTTGAAGAAATGCGTATGACCTTGTGGAAATCCAGATACAAAACTTTCGATGAGCTAT	1036
Db	1006	TGATTTGAAGAAATGCGTATGACCTTGAAGAACTCAAGATATAGAAACTTTGATGAGCTTT	1066
Oy	1031	ATCTCTATTTGTTACTATATGTTGCTGTGACTCTGTAGATTTGATGAGTGTCCAGTATGGGTA	1096
Db	1066	ACCTCTATTTGTTATTAACGTTGTCTGATACGGTTGGGTATATGATGTTCCAAATTAAGGTA	1126
Oy	1091	TTGACCTCGAATCAAAAGGCAACAAGAGAGTATATATATGCTTGGTTTAAAGGC	1156
Db	1126	TTGACCTCGAATCAAAAGGCAACAAGAGAGCTATATATGACCTTTGGCTTTAGGAA	1186
Oy	1151	TTGCAATCAACTTAACCAATATACTCAGAGATGTAGAGAGAGANTCCAGAGAGAGANG	1210
Db	1186	TCGGAAATCAACTTAACCAATATCTCAGAGAGTGTGAGAAAGATCCAGAAAGAGAGAG	1246
Oy	1211	TATACTTGCTTCAGATGATATTAGCAACAGCGGCGCTCCGAGAAAGCATATTTGCTG	1270
Db	1246	TCTACTTACTCTCAAGATGATATTAGCAACAGCGGCGCTCTCTCGACATATACATATTTGCTG	1306
Oy	1271	GAAAGTATCTGATATAGTGAAGGAACTTTATGCAAAACAAATTGAGAGGGCGAGAGAAAT	1336
Db	1306	GAAAGTATCTGATATAGTGAAGAACTTTATGCAAAACAAATTCAGAGGGCGAAGAAAT	1366
Oy	1331	TCTTTGATGAGTCAGAGAAAGGTGTACAGAACTGGACTGTGCTATGATAGGCTGTGT	1396
Db	1366	TCTTGTGATGAGGCAAGAGAGGAGTTACACAACATGAGCTCAGCTTACAGATGCGCTGTAT	1422
Oy	1391	TAAACGCCCTGCTGTTGTATGCAAGATATTTGACGAGATTTGAAGCCACACGATCAACA	1456
Db	1426	GGGCATCTTTGTGTTGTATCCGCCCAAAATATCTGAGAGATTTGAACCCATATGCTCAACA	1486
Oy	1451	ACTTCAACAGAGGGCTTATGTTACCAAGCCAAAGAGCTTCTCACTTGGCCCATTTGCTT	1510
Db	1486	ACTTCAACAGAGACTTATGTGACCAAAACAAAGCTAATTTCTTAACTATTTGCTT	1546
Oy	1511	ATGCAAAATCTTTTGCCCCCTTAATAGAACT-TCCTCTCCACTGACAAAGACATGAATG	1566
Db	1546	ATGCAAAATCTTTTGCCCCCTTAACAGAACTCTTGTCACTCTGTATAGGACATAGACA	1606
Oy	1570	AAGTATGAGTCAATG-----AGTATATACATAAGAACTCAGAGTACTTGTAAA	1622
Db	1606	TCAGATTTAATTAAGCAAGAAACATATATTAATTAATTAAGAAAGAAATTTCTAAA	1666
Oy	1623	TGAGATATCTTT-----TGTAAATGTATCATCAAAAGTAAATGT- AAATTCAT	1674

Db	1666	GTAGTATATGTTGTAATGAGCCACTGTATATCATCAAAAGTAGAGTAATTCAT	1725
QY	1675	ATGACATCTCTTGGTAGAATATTTTCCACTATC	1713
Db	1726	ATAACAATCTCTAGTAGTGTATGTTCACAATCTTAAGC	1764
RESULT 7			
AA093323			
ID	AA093323	standard; cDNA; 1591 BP.	
XX			
AC	AA093323;		
XX			
DT	13-APR-1996	(first entry)	
XX			
DE	Melon phytoene-synthase gene.		
XX			
KM	melon; phytoene-synthase; ripening; cDNA library; fruit; MEL5;		
KM	tomato; TOM5; probe; hybridisation; polymerase chain reaction; PCR;		
KM	antisense; transgenic plant; crop improvement; carotenoid; vector;		
KM	ss.		
XX			
OS	Cucumis melo.		
XX			
PN	MO9602650-A2.		
XX			
PD	01-FEB-1996.		
XX			
PF	06-JUL-1995; 95WO-GB01603.		
XX			
PR	22-SEP-1994; 94GB-0019081.		
PR	18-JUL-1994; 94GB-0014505.		
XX			
PA	(ZENEC) ZENECA LTD.		
XX			
PI	Grierson D, John I, Karvouni Z, Taylor J, Turner A;		
PI	Watson C;		
XX			
DR	WPI; 1996-105912/11.		
XX			
PT	New isolated DNA encoding melon phytoene synthase - used to		
PT	transform plants to modify carotenoid content and related		
PT	characteristics in plant parts, partic. fruit		
XX			
XX	Claim 2; Page 15-16; 22pp; English.		
PS			
XX			
CC	The sequence encodes melon phytoene-synthase (MEL5 gene), and is		
CC	almost full-length. The sequence has been isolated as a cDNA clone		
CC	from a ripening-related cDNA library derived from climacteric melon		
CC	fruit, using the tomato phytoene-synthase cDNA (TOM5) as a		
CC	heterologous probe. The MEL5 gene 5'-end has also been isolated by		
CC	polymerase chain reaction and sequenced. The DNA may be used in		
CC	senes or antisense constructs to modify gene expression in plants.		
CC	The carotenoid content and related characteristics of plant parts		
CC	(particularly fruit) may be modified in this way.		
XX			
XX			
SQ	Sequence 1591 BP; 507 A; 237 C; 395 G; 452 T; 0 other;		
	Query Match	49.1%; Score 847.6; DB 17; Length 1591;	
	Best Local Similarity	77.6%; Pred. No. 1.1e-231;	
	Matches 1136; Conservative	0; Mismatches 249; Indels 79; Gaps 6;	
QY	101	TATATCGTTGAATTAAGTGTAGACTCTAGTGGAATATCTAC--AAGTATGGTTTT	157
Db	116	TAAATTTGTTGAGAGTGGAATAATTTCTTAGTGGGATCTAGAGTAATTTAATTT	175
QY	158	TGATPAAATAGCGCTGAGGTGAGAGTACATAAAGGAAACAAAACCTTGGGAATGT	217
Db	176	CTATPAACTAAGTAAAGTTGGAGGTCACA-AAAAGAAACAAAATCTTGGAATGT	234
QY	218	TTTAGACACCGAGGTTTCTGTTTCAATGAGCAGTGTGTTGTGTGGGTTGTT	277
Db	235	TTTAGACACCAAGT--TTTCTTGCTCAGAAATGCTGTGCTGTATATGGTTT	291



158 TGATTAATAGGCTGAGTGAAGATTAACATTAAGAAAGCAAAAATTGGGAATGT 217  
 113 CTATTAACCTAAGTAAGTTGGAAGTGACA-AAAAGAAAGCAAAAATCTTGGAAATGT 171  
 218 TTATGACCAACGAGGTTCTTGTTCATGAGCATGTCTGTGCTTGTGTGGTGTGT 277  
 172 TTATGACCAACCAAGGT- - -TTCTGTCTCAGATCTGTGCTTGTATGAGGTGT 228  
 278 CTCACCTTCGAGGCTCGAATGGGACAGATTTGTGATTCAGTCCGAGAGGAAC 337  
 229 CTC- - -CTTGTACCTTCGAATGGGACAGATTTGATGAAACATCCGAGAGGAAC 285  
 338 GCGCTTTGTATCATCCAGGTTCTAGCTCGAGATGAGAAATTTGATGGAATGGG 397  
 286 GTTTTGTGATTCATCGAG- - - - -CATAAGAAATTTGGTGTCCATGAGAGAA 333  
 398 TCAGAAAGGTGGGAGACAAAGGTGAATTTGGCTCTTTAATGTGATCCAGATAT 457  
 334 TCATAGAGGT- 344  
 458 CATGCTTGGGTGATCAAGAACTGAAGAGAGCACTTCTCTGATACAGTCCAGTTGG 517  
 345 - 396  
 518 TGGCTAGCCCAAGCTGGAAGATGACTGTGTCATCAGAGAAAAGGTGATGATGTAT 577  
 397 TGGCTACTCCATCTGGAAGACGAGATGACATCGGAACAGATGTCTATGATGTGTT 456  
 578 TAAACGAGGAGCTTTAGTGAAGAGGAGAGCTGAGATCTACCGATGATTTAGAGTGAAG 637  
 457 TGAGCAGAGGAGCCCTTGGTGAAGAGGCAACTGAGATCTACCAATGATGAGTGAAG 516  
 638 CGGATATGTTGTTCAGGGAATTTGGCTGTGTGAGGAAGCATATGATGTTGGCG 697  
 517 CGGATATACCTATTCGAGGGAATTTGGCTGTGTGAGGAAGCATATGATGTTGGCG 576  
 698 AAGTATGAGAGATGAGAAAGCACTTTTCTAGAAACCAAGCTAATGACCCAGAGA 757  
 577 AAGTATGAGAGATGAGAAAGCACTTTTCTAGAAACCAAGCTAATGATGACCCAGAGA 636  
 758 GAAAGAGAGCTATCTGGGCAATATATGTGTGTGAGAGAAAGGATGAGCTTGTGATG 817  
 637 GAAAGAGGCTATCTGGTCAATATATGTGTGTGAGAGAAAGGATGAGCTTGTGATG 696  
 818 GCCCTAATGATCCCAATTAATCTCCGCAAGCTTTAGATGAGTGGAGACAGGCTGAGAG 877  
 697 GCCCAAGGCAATATATTAATCCCGGAGCCTTGTGATGAGTGGAGAAATGAGCTGAGAG 756  
 878 ATATTTTCAAGTGGGCGGCAATTTGATGATGCTGATGCTTATCCGATGATGCTGCA 937  
 757 ATGTTTCAATGGGCGGCAATTTGATGATGCTGATGCTTATCCGATGATGCTGCA 816  
 938 GATTTCTGTTGATATTCAGCCATTCAAGATATGATGAGAAATGGCTATGAGCTTGT 997  
 817 ACTTTCCAGTTGATATTCAGCCATTCAAGATATGATGAGAAATGGCTATGAGCTTGA 876  
 998 GGAATTCAGATACAAAATCTTGCATGAGCTATATCTCTATTTGATGATGCTGCTGTA 1057  
 877 GAAATTCAGATACAAAATCTTGCATGAGCTATATCTCTATTTGATGATGCTGCTGTA 936  
 1058 CTGTAGGATGATGAGTGTTCAGATTATGGGATTTGACCTGATCAAAAGCAACAG 1117  
 937 CGGTTGGGTTGATGAGTGTTCGAATTATGGGATGAGCCCTTGAATCAAAAGCAACAG 996  
 1118 AGAGTGTATATATCTGCTTTGGCTTTAGGGCTTGCAAATCACTAAACCATATATCTCA 1177  
 997 AGAGGCTATATATCTGCTTTGGCTTTGGGATGCGAAATCAATTAATCAATATCTCA 1056  
 1178 GAGATGTAGAGAAATGCGCAAGAGAGAGTATATCTGCTCAAGATGAATTAGCAC 1237  
 1057 GAGATGTAGAGAAATGCGCAAGAGAGAGTATATCTGCTCAAGATGAATTAGCAC 1116

1238 AGGAGGCTCTCCGAGGAAGACATATTTGCTGGAAGTGAAGTGAATGAGAGACT 1297  
 1117 AGGAGGCTATCCGATGAAGATATATTTGCTGGAAGTGAAGTGAATGAGAGACT 1176  
 1298 TTATGAGAAACAAATTCAGAGGCGGAGAAATTTCTTGTATGATGAGAGAGAGTCTCA 1357  
 1177 TTATGAGAAACAAATTCAGAGGCGGAGAAATTTCTTGTATGAGAGAGAGAGTCTCA 1236  
 1358 CAGAACTGAGCTCTGATGATGATGAGTGGCTGTGTTACAGCGCTGCTGTATGCAAGA 1417  
 1237 CAGAACTGAGCTCTGATGATGATGAGTGGCTGTGTTACAGCGCTGCTGTATGCAAGA 1236  
 1418 TATTTGAGAGATTTGAAGCCAGCACTACAACTTTCACAGAGAGGCTTATGTTAGCA 1477  
 1297 TACTAGATGAGATTTGAAGCCAGCACTACAACTTTCACAGAGAGGCTTATGTTAGCA 1356  
 1478 AGCCAAAG- 1536  
 1357 AATCAAGCAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1416  
 1537 AGAATCTCTCTCCACTAGCAAGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1596  
 1417 AACTGCTCTCTCTCAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 1476  
 1597 ACT- 1548  
 1477 AATGATCATTTAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTAT 1536  
 1449 CATCAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1704  
 1537 CATTAATATTTAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1596  
 1705 AACTCATCAAAAGCTCAAGTGAAG 1728  
 1597 ACTTATCTCAACTCTTGAAG 1620

## RESULT 9

AAV03879  
ID AAV03879 standard; cDNA; 1316 BP.

AAV03879;

29-APR-1998 (first entry)

Phytoene synthase coding sequence from *N. benthamiana*.

Phytoene synthase; transgenic plant; enhanced carotenoid synthesis;

ultra violet absorber; food colour; ss.

*Nicotiana benthamiana*.

Key Location/Qualifiers

FT 1..1242

FT /\*tag= a

US5705624-A.

06-JAN-1998.

27-DEC-1995; 95US-0579667.

27-DEC-1995; 95US-0579667.

(DELL/) DELLA-CIOPPA G R.

(FITZ/) FITZMAURICE W P.

(GRIL/) GRILL L K.

(HELL/) HELLMANN G M.

(KUMA/) KUMAGAI M H.

Delia-Cioppa GR, Fitzmaurice WP, Grill LK, Hellmann GM;

Kumagai MH;

XX



Query Match	Best Local Similarity	Score	DB	Length
Matches 1039; Conservative	80.2%;	Pred. No. 2.5e-221;	Indels 84;	Gaps 4
Claim 1; Column 21-26; 25pp; English.				
This sequence encodes the phytoene synthetase from <i>Nicotiana benthamiana</i> . The phytoene synthetase coding sequence represents a cDNA of the invention. The isolated nucleic acid molecules are used for producing recombinant polypeptides or transgenic plants with enhanced ability to synthesise carotenoids. Phytoene has been used as a ultra violet absorber and other carotenoids have been used as food colours, animal feeds and in the pharmaceutical and cosmetics industries.				
Sequence 1316 BP; 403 A; 221 C; 342 G; 350 T; 0 other;				
Query	250	ATGTCGTGCTTTGTTGGGTTTCTCCCACTCGAGAGTCTCGAATGGACAGA	309	
Db	1	ATGTCGTGCTTTGTTGGGTTTCTCCCACTCGAGAGTCTCGAATGGACAGA	57	
Qy	310	TTGTGGATTCAGTCCGAGAAAGAAACCGCGTCTTGTATCATCCAGGTTCTAGCTCGA	369	
Db	58	TTCTTGGATTCAGTCCGAGAGGAAACCGGTTTGTATGGTCGAGG-----	105	
Qy	370	GATAGGATTTGATGATGGAATGGAGGATCAAGAAAGTGGAGACAAAGGTGAAATTT	429	
Db	106	CATAGGAATTTAGTGTGCAATGAGAAACAGAGAGGTGTGGAACAAATGTGGAATTTT	165	
Qy	430	GGCTCTTAAATGCTGATCCAAAGATATTCATGCTTGGGTGATCAAGAACTGAAGAGGA	489	
Db	166	GGTCTCTGTA-----	175	
Qy	490	AGCACTTCTCTGATCCAGTCCAGTTGGTGGCTAGCCAGCTGAGAAAT---GACTGTG	546	
Db	176	-----AGTGTCTATGTTGGGCTACACGGCGGGAGAAATGGCGACGATG	219	
Qy	547	TCATCAGAGAAAGGTGTATGATGTTGTATTAAGCAGCGCTTATGTAAGAGCGAG	606	
Db	220	ACATCAGAAACAGATGTTTATGATGTGATTTGAAACAAACAGCTTATGTAAGAGCGAG	279	
Qy	607	CTGAGATCTACCGATGATTTTGAAGATGGAAGCCGGAATTTGTTCCAGGGAAATTTGGGC	666	
Db	280	TTGAGATCTTACTGATGATTTTGAAGATGGAAGCCGGAATTTGTTCCAGGGAAATTTGGGC	339	
Qy	667	TTGTGATGAGCAATATGATCGTTGTGGCGAAGTATGTGCAAGTATGCAAGACATTT	726	
Db	340	TTGTGATGAGCAATATGATCGTTGTGGCGAAGTATGTGCAAGTATGCAAGACATTT	399	
Qy	727	TACTTAGGAACCAAGCTATATGACCCCAAGAGAGAAGAGCTATCTGGGCAATATATGTG	786	
Db	400	TACTTAGGAACCAAGCTATATGACCCCAAGAGAGAAGAGCTATCTGGGCAATATATGTG	459	
Qy	787	TGGTGCAGGAGAAAGGATGAGTGTGTATGGCCCTATATGCAATCCCAATATCTCCGAA	846	
Db	460	TGGTGCAGGAGAAAGGATGAGTGTGTATGGCCCTATATGCAATCCCAATATCTCCGAA	519	
Qy	847	GCTTATGATAGGTGGAGACAGGCTGGAAGATATTTTCACTGAGGCGGCATTTGATATG	906	
Db	520	GCTTATGATAGGTGGAGACAGGCTGGAAGATATTTTCACTGAGGCGGCATTTGATATG	579	
Qy	907	CTTGATGCTGCTTATCCGATCTGTCTCCAGATTTCTGTTGATATTCAGCCATTTACA	966	
Db	580	CTTGATGCTGCTTATCCGATCTGTCTCCAGATTTCTGTTGATATTCAGCCATTTACA	639	
Qy	967	GATATGATTTGAAGAAATGGTATGGACTTGGGAAATCCAGTAACTTTGGATAG	1022	
Db	640	GATATGATTTGAAGAAATGGTATGGACTTGGGAAATCCAGTAACTTTGGATAG	699	

Oy		1027	CTATATCTCAATTGTACTGTAAGTGGTCAGTACCTGATGAATGAGAGTGTTCCAGTTATG	1068
Db		700	CTAATCCTAATATTGTATTATACGTTGCTGTCGRACAGTTGGCTTAGATGTTCCAATTATG	759
Oy		1087	GGAATTCACCCTGAATCAAAGGCACAACAAGAAGTGTATATAAATGCTGCTTTGGCTTTA	1146
Db		760	GGCATGCCACCTCGAATCAAAGGCCAACAAACAGAGAGTGTATATATATGACGCTTGGCTTGG	819
Oy		1147	GGGCTTCGAATATCACTTAACCAATATTACTCAAGAGTGTAGAGAAAGATGCCAAGAAGGA	1206
Db		820	GGTATCGCGAATCAACTACCAACATTTCTCAGAGATGTCCGAAGAAAGATGCCAAGAAGGA	879
Oy		1207	AGAGTATATCTTGCTCCCAAGATGAATTAGCAACAGCAGAGGCTCTCCGACGAACATATTT	1266
Db		880	AGAGTCTACTTACTCTCAAGATGAATTAGCACAGCAGAGTCTCTCCGACGATGACATATTTT	939
Oy		1267	GCTGGAAGAGTGACTGTATAGTGGAGAACCTTTATGAAGAAACAAATTCAGAGGCGAGG	1326
Db		940	ACTGGAAGAAGTGACTGTATATAGGAAGCTTTATGAAGAAGCAATTCAGAGGCGAAGA	999
Oy		1327	AAATTCCTTTGATGAGTCAGAGAAAGGTGTACAGAACTGCACTTGCTAGTAGATGGCT	1386
Db		1000	AAGTCTTTCATAGAGGAGAGGAAGAGATTACACAACTGAGCTTCAGCTAGCAGATGGCCT	1059
Oy		1387	GTTGTTAACAGGGCGCTGTTGTATCGAAGATATTGGACGAGATTGAAGCCAACGACTAC	1446
Db		1060	GTAAGGGGACTTTTGCTGTGTGTACCGCCAATTAICTGACAGAGATCGAAGCCAAATGACTAC	1119
Oy		1447	AACAACCTTCACAAGAGAGGCTTATGTATTAGCAAGCCAAAAGACTTCTCACCTTGCCCAAT	1506
Db		1120	AACAACCTTCACAAGAAGAGCTTATGTATGAGCAAAATCAAAGAAGCTAATTTCTTACCTATT	1179
Oy		1507	GCTTATGCAAAATCTCTTGTCGCCCTTAATAGAACT	1542
Db		1180	GCTTATGCAAAATCTCTTGTCGCCCTTAACAAGAACT	1215
 RESULT 10 AAZ99482 standard; cDNA; 1239 BP.				
XX	AA	AAZ99482;		
AC				
XX				
DT	03-JUL-2000	(first entry)		
DE	cDNA encoding a phytoene synthase polypeptide.			
XX				
KW	Gibberellin acid; copalyl diphosphate synthase; 3beta-hydroxylase;			
KM	2-oxidase; phytoene synthase; C-20 oxidase; 2beta,3beta-hydroxylase;			
KW	seed germination; seedling growth; gibberellin biosynthetic pathway;			
KM	transgenic plant; hypocotyl; epicotyl; ss.			
XX				
OS	Lycopersicon esculentum.			
XX				
FH	Key	Location/Qualifiers		
FT	CDS	1..1239		
FT		/**tag= a		
FT		/product= "phytoene synthase"		
FT		/transl_except= (pos: 1027..1029, aa: Lys)		
FT		/transl_except= (pos: 1057..1059, aa: Arg)		
XX				
PN	MO200009722-A2.			
XX				
PD	24-FEB-2000.			
XX				
PF	10-AUG-1999;	99WO-US18066.		
XX				
PR	10-AUG-1998;	98US-0096111.		
PR	07-JUN-1999;	99US-0137977.		
XX				
RA	(MONS ) MONSANTO CO.			
XX				



DR MPI, 2002-489107/52.  
 DR P-PSDB; AAE24923.  
 XX Control of gibberellin levels in plants useful to avoid unfavorable  
 PT conditions in crops to increase yields, using transgenic plants having  
 PT reduced seed germination and early seedling growth then treatment to  
 PT restore these properties -  
 XX  
 XX Claim 45; Page 99; 155pp; English.  
 XX  
 CC The invention relates to control of gibberellin (GA) levels in plants.  
 CC The method involves producing transgenic plants having a phenotype  
 CC of reduced seed germination and reduced early seedling growth, then  
 CC restoring seed germination and early seedling growth by treating  
 CC plants with an appropriate compound when conditions are favourable.  
 CC The method is useful to control seed germination and/or early seedling  
 CC growth in agricultural production so that unfavorable environmental  
 CC conditions normally reducing agronomic output can be avoided and  
 CC yields increased. Plants also demonstrate increased uniformity of  
 CC germination, emergence and seedling vigor, so increasing yields at  
 CC harvest. The method is especially useful in crop plants such as e.g.  
 CC canola, soybean, cotton, etc., and is also useful in storage and  
 CC transport of seeds to reduce premature germination which may affect  
 CC agronomic or food quality of the seeds. The present sequence is  
 CC tomato phytoene synthase cDNA. This cDNA is used in exemplification  
 CC of the invention.  
 CC  
 XX  
 XX Sequence 1239 BP; 376 A; 208 C; 316 G; 339 T; 0 other;

Query Match 46.7%; Score 806.4; DB 24; Length 1239;  
 Best Local Similarity 83.5%; Pred. No. 5.7e-220;  
 Matches 914; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 466 GGTGATCAAGAACGAAAGGAGCACTTCTCTGACGTCAGTTGGTGGTAC 525  
 DB 145 GGTGAAAGCAACTAATATATGACGAAATTTCTGTACGGCTGCTATTTTGGCTACT 204  
 QY 526 CCAGCTGAGAAATGCTGTCTATCAGAGAAAAGGTATGATGATTTAAAGCAG 585  
 DB 205 CATCTGAGAAAGCAATGACATGCAACAGATGCTATGATGATGATTTGAGGACG 264  
 QY 586 GCAGCTTTAGTAAAGGAGCGAGCTGAGATCTACGATATTGAAAGCGGATATT 645  
 DB 265 GCAGCTTGTGAAAGGAGCACTGAGATCTACCAATGAGTTGAAGTGAAGCGGATATA 324  
 QY 646 GTTGTCCAGGAATTTGGGCTTTGATGAGCAATGATGCTTTGGCCGAAATATGT 705  
 DB 325 CCTATTCGGGAAATTTGGGCTTTGATGAGCAATGATGAGTGGTGGTGAAGTATGT 384  
 QY 706 GCAGATATGCAAAAGCATTTTACTTAGAACCAAGCTAATGACCCGAGAGAAAGAGA 765  
 DB 385 GCAGATATGCAAAAGCGTTTAACTTAGAACATGCTAATGCTCCGAGAGAAAGAG 444  
 QY 766 GCTATCTGGGCAATATATGATGTCAGAGAAACGATGAGCTTTGATGGCCCTAAT 825  
 DB 445 GCTATCTGGGCAATATATGATGTCAGAGAAACGATGAGCTTTGATGGCCCAAC 504  
 QY 826 GCATCCCAATATCTCCGAGCTTTAGATAGTGGGAAACAGGCTGGAAGATATTTTC 885  
 DB 505 GCATCAATATATATCCCGGAGCCCTAGATAGTGGGAAATAGGCTAGAAATGTTTC 564  
 QY 886 AGTGGGCGGCAATTTGATGCTGATGCTTTATCCGATACGTCCTCCAGATTTTCT 945  
 DB 565 AATGGGCGGCAATTTGATGCTGATGCTTTATCCGATACGTCCTTTAACTTTTCA 624  
 QY 946 GTTGATATTCAGCCATTGAGATATGATGAAAGAAATCGATGAGCTTTGGAATCC 1005  
 DB 625 GTTGATATTCAGCCATTGAGATATGATGAAAGAAATCGATGAGCTTTGGAATCC 684  
 QY 1006 AATATCAAACTTTGATGAGCTATATCTCTATTTGATCTAATGTTGGTGTCTGTAGA 1065  
 DB 685 AATATCAAACTTTGATGAGCTATATCTCTATTTGATCTAATGTTGGTGTCTGTAGA 744

QY 1066 TTGATAGTGTCCAGTTATGGTATTTGACCTGTAATCAAGGCAACAGAGAGTGA 1125  
 DB 745 TTGATAGTGTTCATTTATGGGTATTCGCTGATCAAGGCAACAGAGAGCGTA 804  
 QY 1126 TATTAATGCTGCTTTGGCTTTAGGCTTTGCAATCAATCAATATATCTGAGATGTA 1185  
 DB 805 TATTAATGCTGCTTTGGCTTTAGGCTTTGCAATCAATCAATATATCTGAGATGTA 864  
 QY 1186 GGAGAAAGATGCCAGAAAGGAAAGATATATCTTCCCTCAAGATGATTTGCAACAGCAGG 1245  
 DB 865 GGAGAAAGATGCCAGAAAGGAAAGATATATCTTCCCTCAAGATGATTTGCAACAGCAGG 924  
 QY 1246 CTCTCCGAGCAAGATATTTGCTGGAAGAGTGAATAGTGAAGCACTTTATGAAG 1305  
 DB 925 CTATCGATGAAAGATATATTTGCTGGAAGAGTGAATAGTGAAGCACTTTATGAAG 984  
 QY 1306 AAACAATTCAGAGGCGGAGAAATCTTTGATGATGATGATGATGATGATGATGATG 1365  
 DB 985 AAACAATTCAGAGGCGGAGAAATCTTTGATGATGATGATGATGATGATGATGATG 1044  
 QY 1366 GACTGTGCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1425  
 DB 1045 AGCTAGCTAGTATATATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1104  
 QY 1426 GAGATTGAAGCCAGCAATCAACCACTTCAACAGAGGCTTATGATGATGATGATGATG 1485  
 DB 1105 GAGATTGAAGCCAGCAATCAACCACTTCAACAGAGGCTTATGATGATGATGATGATG 1164  
 QY 1486 AAGCTTTCACCTTCCCATTTGCTTATGCAAAATCTTTGCTCCCTTAATGAACTTCC 1545  
 DB 1165 AAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1224  
 QY 1546 TCTCCATGAGCAA 1559  
 DB 1225 TCTCTCAAGATA 1238

RESULT 12  
 AA16951  
 ID AA16951 standard; cDNA to mRNA; 2868 BP.  
 XX  
 XX AA16951;  
 AC XX  
 XX  
 DT 06-JUL-1998 (first entry)  
 XX  
 DE Nucleic acid encoding phytoene synthase 4.  
 XX  
 KW Phytoene synthase; breeding; variable flower colour; ds.  
 XX  
 OS Gentiana lutea.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 400..1689  
 FT /\*tag= a  
 PN JPI0084966-A.  
 XX  
 PD 07-APR-1998.  
 XX  
 PF 17-SEP-1996; 96JP-0245107.  
 XX  
 PR 17-SEP-1996; 96JP-0245107.  
 XX  
 PA (IWAT-) IWATE KEN.  
 XX  
 DR MPI; 1998-264853/24.  
 DR P-PSDB; AAW46964.  
 XX  
 PT Phytoene synthase gene - useful for breeding plant of variable  
 PT flower colour  
 XX  
 PS Claim 4; Pages 12-14; 15pp; Japanese.





```

Db      636 TGGCAATGTTTTCTTGAGCCAAATTCGAAAAATTACATTTTTG----- 681
Oy      360 CCTAGCTGAGATAGAAATTTGATGTGGAAATGGAGAAATCAAGAAAGTGAGACAAG 419
Db      682 -----GATAAAGTTTATGTATACAAATGGAAGTTAAGAAAAGTACACCAAG 731
Oy      420 GTGGAATTTGGCTTTTAATTTGCTGATCCAAGATATTCATGCTGGGTGATCAAGAC 479
Db      732 GCGTAGATCAGGTATGAGGGTGTGAGATTTGAGTTCATTTTGTCTTGAGAGAGTCTGATT 791
Oy      480 TGA---AGGGAAGCATTTCTCTGTACAGTCCAGTTTGGGTGATGCCAGCTGAGAGA 536
Db      792 AGAGACCCCGGGAAGAAATTAATCGGGTATCCTCATATTTAAGTTAACCCCGCAGAGAG 851
Oy      537 AATGATGTGTATCAGAGAAAAAGGTGATGATGTGGATTTAAGCAGCAGCTTTAGT 596
Db      852 AAATGACATGACATAGAGCAAAAGTTTATGATGTCTTTTAAAGCAAGCAGCTTTGAT 911
Oy      597 GAAGAGCAGCTGAGATCTACCGATGATTTAGAACTGAAGCCGATATTTGTTCCAGG 656
Db      912 TAATAGACAGTTGAGGTCTAGAGAAATTTGAGAGTGAACCGGACATTTATTTGCCAGG 971
Oy      657 GAATTTGGGCTTTGATGAGCAATATGATGTTGTGCGCAAGATATGTGACAGATATGC 716
Db      972 AAACCGCAACGTTGAATGAAGCTTATGATCGGTGCGAGAAATGTGCTGAATATGC 1031
Oy      717 AAAGCATTTTACTTAGGAACCAAGCTATATGACCCGAGAGAAAGAGAGCTATCTGGGC 776
Db      1032 CAAGCATTTCTACTGGGGAACCCAGCTATGACACCGAGAGGCTTTAGCTATCTGGGC 1091
Oy      777 ---AATATATGTGTGTGCAAGAGAACGATGAGCTTTGATGAGCCCTAATGATCCCA 833
Db      1092 CGATATATATGATGTGCTAGAGAGACAGATGACCTTTGATGAGGCTTAACGCTGACA 1151
Oy      834 CATACTCCGCAAGCTTATGATGAGTGGAGAACCGAGCTGAGAAATTTTCAAGTGGCG 893
Db      1152 CATAAATCACCCGCTTGAATGATGAGTGGAGCAAGATCAGAAAGTGTTCACAGGCA 1211
Oy      894 GCCATTTGATATGCTTGTGATGCTGCTTTATCCGATCTGCTCCAGATTTCTGTTGATAT 953
Db      1212 ACCTCTGATATGCTTGTGATGCTGCTTTATCTGATACCATTTACAGTATCTGTGACAT 1271
Oy      954 TCAGCATTCAAGATATGATGAAAGATGCGTATGGAATTGGAATCCAGATACAA 1013
Db      1272 CGAGCATTTAGATGATGATGAAAGATGCGATGATGCAAGAAATCGAATACAA 1331
Oy      1014 AACTTTCATGAGCTATATCTTATTTGATCTATGTTGCTGTACTGTAGATTTGATGAG 1073
Db      1332 GAATTTTCATGAGCTGTATCTTATCTGATTTATGTGCTGTGACACAGTTGGCTGTATGAG 1391
Oy      1074 TGTTCAGTTATGAGGTATTTGACCTGAAATCAAAAGGCAACAAGAGAGTGTATATATGC 1133
Db      1392 TGTACCAATATGAGCATTTGACCTGAATCTTAAGCAACACAGAAAGTGTGTATGATGC 1451
Oy      1134 TGCCTTGGCTTTAGGGCTTGAATCAATCAATCAATATCTCAGAGATGTAGAGAAAG 1193
Db      1452 AGCTTATCTTTGGGATCGCGAACACAGCTGACTAATCTTAAGAGATGTTGAGAAAG 1511
Oy      1134 TCCCAAGAGAGAAAGATATCTTGTCTTCAAGATGAATTAAGACAGGAGCTCCGA 1253
Db      1512 TGCAGAAAGAGAGAGATGTATCTTACCTTCAGATTAATTAAGCAAGCAGGTTTATCGA 1571
Oy      1254 CGAAGACATTTATTTGCTGGAAGAGTGAATGAATGAGAGAACTTTATGAGAAACAAT 1313
Db      1572 TGAGACATTTTGTCTGAAAAGTTTACAGACAAATGAGAGATTTTATGAGAAACAAT 1631
Oy      1314 TCAGAGGCGAGAGAAATTTCTTTGATGAGTCAAGAGAAAGTGTCAAGAACTGAGCTTGC 1373
Db      1632 CAAAAGGGGTAGAAAATTTCTATGATGAGCAAGAAAGAGTCCCGAAGTCAAGTCCGC 1691
Oy      1374 TAGTATGATGAGCTGTGTTTAAACAGCGCTGTGTTGATGCAAGATTTGAGCAGATTTGA 1433
Db      1692 GAGCAGATTTGCTGTGTGGGAGCGTGTCTTTTATAGAAAAAATTTGATGAGATGAGA 1751

```

```

Oy      1434 AGCCAAAGACTATCAACAATCTTCAACAGAGGCTTATGTTAGCAAGCCAAAGACTTCT 1493
Db      1752 AGCAATATGACTACAAACAATTTCCGCCAGAGGGCTTATGTTAAGAGCGAAGAACTATT 1811
Oy      1494 CACCTGGCCATTTGCTTATGCAAAATCTCT 1523
Db      1812 AGCTATGCTGTATGATGATGCGCAAGTCTCT 1841

RESULT 15
AAV16950
ID AAV16950 standard; cDNA to mRNA, 1932 BP.
AC AAV16950;
XX
AC AAV16950;
XX
DT 06-JUL-1998 (first entry)
XX
DE Nucleic acid encoding phytoene synthase 3.
XX
KW Phytoene synthase; breeding; variable flower colour; ds.
XX
OS Gentiana lutea.
XX
FH Key Location/Qualifiers
FT CDS 499..1785
FT /tag= a
XX
PM JPI0084966-A.
XX
PD 07-APR-1998.
XX
PF 17-SEP-1996; 96JP-0245107.
XX
PR 17-SEP-1996; 96JP-0245107.
XX
PA (IWATE-) IWATE KEN.
XX
XX WPI; 1998-264853/24.
XX
DR P-PSDB; AAW46963.
XX
PT Phytoene synthase gene - useful for breeding plant of variable
PS flower colour
XX
PS Claim 3; Pages 10-12; 15pp; Japanese.
XX
CC The present sequence encodes phytoene synthase 3. It was isolated from
CC a cDNA library prepared from mRNA extracted from the petals of Gentiana
CC lutea. The nucleic acid sequence was amplified from the library using
CC PCR primers AAV16952-53. The phytoene synthase gene is useful for
CC breeding plants with variable flower colours.
XX
SQ Sequence 1932 BP; 635 A; 302 C; 434 G; 561 T; 0 other.

Query Match 37.8%; Score 653.4; DB 19; Length 1932;
Beet Local Similarity 72.6%; Pred. No. 4e-176;
Matches 940; Conservative 0; Mismatches 321; Indels 34; Gaps 6;

Oy      240 TTTTATGAGCATGCTGTGCTTTGTTGTGTGCTGTTTCTCCCACTCCGAGTCTGAA 299
Db      489 TTTGTTAAACATGCTTATTTGATGCTATGAGTGTGCTGCGAGTCTTGAAATTTTGA 548
Oy      300 TGGGACAGATTTGTTGATTCAGTCCGAGAGAAACCGGCTTTTATCATCCAGGT 359
Db      549 TGGCAATGTTTTCTTGAAGCCAAATTCAGAAAGTTACATTTTTCG----- 594
Oy      360 CCTAGCTCAGATAGAAATTTGATGTGAATGGAGATGGAATCAAGAAAGTGGAGACAAG 419
Db      595 -----GATAAAGTTTAAATGTAACAATGGAAGATTAAGAAAGTAGACACCAAG 644
Oy      420 GTGGAATTTGGCTTTTAAATTTGCTGATCCAAGATATTCATGCTTGGGTGATCAAGAAC 479
Db      645 GCGTAGATCAAGTTATGAGGGTGTGAGATTTGAGTTCAATTTGCTTGAAGAGTCTGATT 704

```

QY 480 TGAAG--AGGAGACATTTCTCTGTACAGTCCAGTTTGCTAGCCAGCTGAGAG 536  
DB 705 AGAGACCCCGGAGAGAAATATCGTATCTCCAGTATATAGCTACCCCGGAGAGAG 764  
QY 537 AATGATGTGTATCAGAGAAAAGGTGATATGTGTATTAAGCAGGCGCTTTAGT 596  
DB 765 AATGACGATGATCAGAGCAAAAGTTATGATGTCGTTTAAAGCAAGCAGCTTTGAT 824  
QY 597 GAAGAGCAGCTGAGATCTACCGATGATTTAGAGGAAAGCCGATATGTTGTTCCAG 656  
DB 825 TATATGACAGTTGAGCTCTAGAGAAAATTTGAGGTAAGCCGACATTTATTTGCCAG 884  
QY 657 GAATTTGGGCTTTGAGTGAAGCATATGATGTTGGCGAAGTATGTGACAGATATGC 716  
DB 885 AAACCGAAGCGTTGAAATGAAGCTTATGATCGGTGTGAGAAATGATGTGTAATATGC 944  
QY 717 AAAGCATTTTACTTAGAAGCAAGCTAATGACCCGAGAGAGAAAGCTATCTGGGC 776  
DB 945 CAAGTCATTCTAAGTGGGAAACCACTCATGACCCGAGAGCGCTTTAGCTATCTGGGC 1004  
QY 777 AATATATGTGTGACAGAGAAACGATGAGCTTGTGATGGCCCTAATGATCCACAT 836  
DB 1005 GATATATGTATGTGTAGAGAGACAGATGAGCTTGTGATGGCTTAAGCGCTCAAAAT 1064  
QY 837 AACTCCGACAGCTTTAGATAGGTGGGAGACAGGCTGAAAGATATTTTCAGTGGCGGCC 896  
DB 1065 AAATCCAAACCGCTTATGATAGGTGGGAAAGCAAGATTAAGAGATGTTTCAAAAGGCAACC 1124  
QY 897 AATTGATATGCTTTGATGCTGCTTTATCCGATACGTCTCCAGATTTCTGTTGATATTCA 956  
DB 1125 TTTTATATATGCTTGTATGCTGCTTTATCTGATACCAATTAACAGTATCTGTGACATCCA 1184  
QY 957 GCCATTGAGATATGATTTGAAGAAATGCGTATGAGCTTGTGAATCCAGATACAAAC 1016  
DB 1185 GCCATTGAGATATGATTTGAAGAAATGCGTATGAGCTTGTGAATCCAGATACAAAC 1244  
QY 1017 TTTTCATGAGCTATATCTCTATTTACTATGTGTGCTGCTGCTGATGATGATGT 1076  
DB 1245 TTTTCATGAGCTATATCTCTATTTACTATGTGTGCTGCTGCTGATGATGATGT 1304  
QY 1077 TCCAGTATGGGTATGCACTGGAATCAAAGGCAACACAGAGCTATATATATGCTGC 1136  
DB 1305 ACCAGTAAATGGGCATTTGCACTGGAATCTTAAGGCAACACAGAGCTATATATGCAAGC 1364  
QY 1137 TTTGGCTTAAAGGCTTGCATACTCAATCACTAATCTCAGAGATGTAGAGAAATGC 1196  
DB 1365 TTTATCTTTGGGATCGCGAACCACTGCTAACAATTTTAAGGATGTTGGAGAAATGC 1424  
QY 1197 CAGAAGAGAGATATATCTCTCAAGATGAATTAAGCAGAGGCTCTCCGACGA 1256  
DB 1425 AAGAAGAGGAGAGATGATCTTACTCTCAAGATGAATTAAGCAGAGGCTTTATCAAGATGA 1484  
QY 1257 AGACATATTTGCTGGAAGAGTGAATGAAGAGAACTTTATGAAGAAACAATTC 1316  
DB 1485 GGACATATTTTGTGGAAGAGTTCCAGACAAATGAAGATTTTATGAAGAAACAATTC 1544  
QY 1317 GAGGGCGAAGAAATCTTTGATGA-CTCAGAGAAAGGTGTCAAGAACTGGAAGCTG-CT 1374  
DB 1545 AAGGGCTAGAAAATTTATGATGATGCGCAGAAAAGGTCGCCGAACTCAAGCTCCGCG 1604  
QY 1375 AGTAGATGGCTGTGT---TAAACAGCGTGTGTTATCCAGATATTTGACAGAT 1430  
DB 1605 AGCAGATTTGCTGTGTGGGCGAGCGTTTCTTTTATAGAGAAAATTTGATGAGAT 1664  
QY 1431 TGAAGCCAAAGCACTAACCAACTTC-ACAAGAGAGGCTTATGTAAGCAAGCCAAAGAAC 1489  
DB 1665 AGAAGCAATGACTACAACTTCAACAAAGGCTTATGTAAGCAAGCCAAAGAAC 1724  
QY 1490 TTCTCACTTGGCCATGTCTTATGCAAAATCTCTT 1524  
DB 1725 TATTAGCTATGCTGTAGCATGTGCCAAGTCTCTT 1759

Search completed: January 16, 2004, 01:16:29  
Job time : 497 secs







Matches 1187; Conservative 0; Mismatches 255; Indels 96; Gaps 7;

```

OY 109 TTGAATAGTGGATAGACTCTAGTGGATATCTACAGATATGTTTTTGTAAATAG 168
DB 221 TAGACTTAGTGGGAATCTACTAGAGATATTTTTTATTTTTTATTAATTA 280
OY 169 GCTGAGTGGAGGTAACATTA-----GGAAGACAAAACCTGGAAATGTTT 220
DB 281 GCAGAGGAGGAGGAAACAGAAAACAGAAATAGACAAAACCTGGAAATGTTT 340
OY 221 AGACCAAGAGGTTCTTGTTCATGAGCATGTCTGTGCTTTGTGGGTTGTTCTC 280
DB 341 AGACCAAGAGGTTCTTGTTCATGAGCATGTCTGTGCTTTGTGGGTTGTTCTC 397
OY 281 CCACCTCGAGGTTCTGATGGAGAGATGTTGGATTCAGTCCGAGAGGAAACCGG 340
DB 398 ---CTTGGAGGTTCTCAATGGAGAGATTTTGGATTCATTAAGGAGGAAACCGG 454
OY 341 TCTTGTATCATCCAGGTTCTAGCTCGAGATGAGATTTGATGGAATGGAGATCA 400
DB 455 TTTTGTACTCGTCGAG-----CATAGAAATTTAGTGTGCAATGAGAAATCA 502
OY 401 AGAAGGTGGAGAGACAAAGGTGGAATTTTGGCTCTTTAATGCTGATCCAGATTTAT 460
DB 503 AAGAGGTGGAGAGACAAAGGTGGAATTTTGG-----533
OY 461 GCTTGGGTGGATCAAGAACTGAAAAGGAGACCTTCTGTGATAGTCCAGTTGGTG 520
DB 534 -----TTCGTAGCGGTGCGATGGTG 556
OY 521 CTAGCCAGCTGAGAAAT---GACTGTGTATCATGAGAAAAGGTGATGATGTAT 577
DB 557 CTACACCAACGGGAGAAATGGGCAATGACATCAGAACAGAGTTTATGATGTAT 616
OY 578 TAAAGCAGGAGCTTTAGTGAAGAGGAGCTGAGATCTACGATATTTAGAAAGTGAAC 637
DB 617 TGAACCAAGAGCTTTAGTGAAGAGGAGCTGAGATCTACGATATTTAGAAAGTGAAC 676
OY 638 CGGATATTTGTTGTCAGGGAATTTGGGCTGTGATGGAAGCATATGATGTTGGTG 697
DB 677 CGGATATTTGTTGTCAGGGAATTTGGGCTGTGATGGAAGCATATGATGTTGGTG 736
OY 698 AAGTATGTCAGAGTATGCAAAAGACATTTTACTAGAACCAAGCTTAATGACCCAGAGA 757
DB 737 AAGTATGTCAGAGTATGCAAAAGACATTTTACTAGAACCATATGCTAATGATCCAGAGA 796
OY 758 GAAAGAGCTATCTGGGCAATATATGTGTGTGCAAGAAACGATGAGCTTTGTATG 817
DB 797 GAAAGAGCTATTTGGGCAATATATGTGTGTGCAAGAAACGATGAGCTTTGTATG 856
OY 818 GCCCTAATGCAATCCCAATCTCCGCAAGCTTTAGATAGGAGGAGACGAGGCTGAGAG 877
DB 857 GCCGAAATGCAATCAATATCTCCCAAGCTTTAGATAGGAGGAGACGAGGCTGAGAG 916
OY 878 ATATTTTCAAGTGGGCGCATTTGATGATGCTTGTATGCTTATCCGATACGTCTCCA 937
DB 917 ATGTTTCAAGTGGGCGCATTTGATGATGCTGATGCTTGTGCGATGCTGTTCC 976
OY 938 GATTTCCTGTTGATATTCAGCCATTCAGAGATATGATGGAAGAAATGGTATGACCTGT 997
DB 977 AGTTTCAGTGTATTCAGCCGTTCAAGATATCATTTGAAGAAATGGTATGACCTGTGA 1036
OY 998 GGAATCCAGATACAAACTTTCGATGAGCTATCTATGTTACTATGATGCTGTGTA 1057
DB 1037 GGAAGTCAAGATACAAACTTTCGATGAGCTATCTATGTTACTATGATGCTGTGTA 1096
OY 1058 CTGTAGATGATGAGTGTTCAGATTTAGGATTTGACCTGATCAAGAGCAACAG 1117
DB 1097 CGGTGGGTGATGAGTGTTCAGATTTAGGATTTGACCTGATTTCAAGAGCAACAG 1156
OY 1118 AGAGTATATATGCTGCTTTGGCTTTAGGCTTGAACATCACTAAACAATATCTCA 1177
DB 1157 AGAGTATATATGCTGCTTTGGCTTTAGGCTTGAACATCACTAAACAATATCTCA 1216

```

```

OY 1178 GAGATGAGAGAGATGCCAGAGAGAGATATCTTGCTCAAGATGATTAACAC 1237
DB 1217 GAGATGTCGAGAGAGATGCCAGAGAGAGATCTTACTTACCAAGATGATTAACAC 1276
OY 1238 AGCAGGCTCTCCGAGAGACATATTTGCTGGAAGAGTACTGATTAAGAGAACT 1297
DB 1277 AGCAGGCTCTCCGAGAGATGACATTTTGTCTGGAAGAGTACTGATTAAGAGAACT 1336
OY 1298 TTATGAAGAAACAAATTCAGAGGCGAGAAATCTTGTATGATGAGAGAAAGTGTCA 1357
DB 1337 TTATGAAGAAACAAATTCAGAGGCGAGAAATTTCTGACAGGCGAGAGAGAGTGA 1396
OY 1358 CAGAACTGACTCTGCTAGTATGATGAGCTGTGTTAAGAGGCTGCTGTATGCAAGA 1417
DB 1397 CAGAACTGAGTCTGCTAGTATGATGAGCTGTGTTAAGAGGCTGCTGTATGCAAGA 1456
OY 1418 TATGAGACAGATTTAAGCCAGACTACAACTTCAAGAGAGGCTTATGTTGCA 1477
DB 1457 TACTGACAGATTTAAGCCAGACTACAACTTCAAGAGAGGCTTATGTTGCA 1516
OY 1478 AGCCAAAGAGCTTCTGACCTTCCATGCTTATGCAAAATCTTGTGCCCCCTATA 1537
DB 1517 AACCAAGAGCTATATTTCTTACTTATGCTTATGCAAAATCTTGTGCCCCCTATA 1576
OY 1538 GAATCTCTCTCCACTAGCAAGAGACATGATGAGTGTGATGATTAATTAAC 1596
DB 1577 GAATCTCTCTCCACTAGCAAGAGACATGATGAGTGTGATGATTAATTAAC 1636
OY 1597 ACTAAGAACTCAGTACTTGTATGATGATGATTTT 1634
DB 1637 TATATTTATTTATTTATGTTTAAAGAAAGATTT 1674

```

# RESULT 2

US-08-579-667-1

; Sequence 1, Application us/08579667

; Patent No. 5705624

## GENERAL INFORMATION:

APPLICANT: Fitzmaurice, Wayne P.

APPLICANT: Hellmann, Gary M.

APPLICANT: Grill, Laurence K.

APPLICANT: Kumagai, Monto H.

APPLICANT: Della-Cioppa, Guy R.

TITLE OF INVENTION: DNA SEQUENCES ENCODING ENZYMES USEFUL IN

TITLE OF INVENTION: PHYTOENE BIOSYNTHESIS

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Virginia C. Bennett

STREET: 1211 East Morehead Street, PO Drawer 34009

CITY: Charlotte

STATE: No. 5705624th Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/579,667

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Bennett, Virginia C.

REGISTRATION NUMBER: 37,092

REFERENCE/DOCKET NUMBER: 627-196

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-420-2200

TELEFAX: 919-881-3175

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1795 base pairs





TITLE OF INVENTION: DERIVED THEREFROM  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cushman, Darby & Cushman  
STREET: 1615 L Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20036-5601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/995,950  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/625,664  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kokulis, Paul N.  
REGISTRATION NUMBER: 16,773  
REFERENCE/DOCKET NUMBER: PNK/3893/82895/MJM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3000  
TELEFAX: 202-822-0944  
TELEX: 248453 cusb  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1646 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-07-995-950-2

Query Match 48.7%; Score 842; DB 1; Length 1646;  
Best Local Similarity 74.3%; Pred. No. 4.8e-232;  
Matches 1222; Conservative 0; Mismatches 330; Indels 92; Gaps 9;

QY 101 TTTAAATCGTTGAATTAATGATAGACTCTAGTGATATCTAC---AAGTATGTTT 157  
DB 53 TAAATTTGTTGAGAGTGAATATTTCTTAGTGGAACTCTAGAGATTAATTTATTT 112  
QY 158 TGATAAATAGAGCTGAGTGAAGGTAACTAAGAGAAACAAAACTTGGAAATGT 217  
DB 113 CTATAAACTAAGTAAAGTTTGGAGGTGACA-AAAAGAAACAAAATCTTGGAAATGT 171  
QY 218 TTTAGACCCGAGGTTTCTTGTTCATGAGCATGCTGTTGTTGTTGGTTT 277  
DB 172 TTTAGCAACCAAGT---TTTCTTGCTCAGAAATGCTGTGCTTGTATGGTTT 228  
QY 278 CTCGCCACTTCGAGGCTCGAATGGAGAGATTTGATTGATCCGAGAGAAACC 337  
DB 229 CTC---CTTGTAACGCTCAATAGGACAAAGTTTCAATGATCAGTCGGAGGAAACC 285  
QY 338 GCGTCTTTGATTCATCCAGGTTCTAGCTCGAGATAGAAATTTGATGTAATGGAGAA 397  
DB 286 GTTTTTTTGATTCATGAGG-----CATAGGAATTTGGTGTCCATAGAGAGAA 333  
QY 398 TGAAGAAAGTGGAGAGCAAAAGTGAATTTGGCTCTTAATGCTGATCCAAATAT 457  
DB 334 TCAATAGAGT----- 344  
QY 458 CATGCTGGTGGATCAAGAACTGAAAGGAGACATTTCTGTACAGTCGAGTTGG 517  
DB 345 -----GCTGAAGCAAACTAATATAGAGAGAAATTTCTGTACGCTCTGATAT 396  
QY 518 TGGCTAGCCAGCTGAGTAATGATCTGTCTCATCAGAGAAAAGGTGATGATGTTAT 577  
DB 397 TGGTACTCTCATCTGAGAGAGAGAGATGATCATCGAGACAGATGCTATGATGTT 456

QY 578 TAAAGCAGCAGCTTTAGTGAAGAGAGCTGAGATCTACCGATGATTTAGAGTGAAGC 637  
DB 457 TGAGCAGCAGCCTTTGGTGAAGAGCAATGAGATCTACCAATAGATTTAGAGTGAAGC 516  
QY 638 CGGATATTTGTTGTTCCAGGAAATTTGGCTTTGAGTGAAGCATATGATCGTTGGCG 697  
DB 517 CGGATATACCTATTCGGGGAAATTTGGCTTTGAGTGAAGCATATGATGTTGGTG 576  
QY 698 AAGTATGTCAGAGTATGCAAGACATTTTACTTAGAACCAAGCTAATGACCCAGAGA 757  
DB 577 AAGTATGTCAGAGTATGCAAGACATTTTAACTTAGAACCATATGCTATGATCCTCCGAGA 636  
QY 758 GAAAGAGAGCTATCGGGCAATATATGTTGTTGAGAGAGAACGATAGCTGTTGATG 817  
DB 637 GAAAGAGAGCTATCGGGCAATATATGTTGTTGAGAGAGAACGATATGATGTTGATG 696  
QY 818 GCCCTAATGATCCCACTAACCTCCGAAAGCTTTAGATAGTGGAGACCAAGCTGGAAG 877  
DB 697 GCCCAAGCATCATATATTACCCGAGCCTTAGATAGTGGGAAATAGGCTAGAG 756  
QY 878 ATATTTTCAATGAGGCGGCAATTTGATATGCTGATGCTGTTTATCCGATATCTGCTCA 937  
DB 757 ATGTTTCAATGAGGCGGCAATTTGATGATGCTGATGCTGTTTATCCGATATCTGCTCA 816  
QY 938 GATTTCTGTTGATTTAGGCAATTCAGAGATGATGATGAAAGAAAGGCTATGAGCTTGT 997  
DB 817 ACTTTCCAGTTGATTTAGGCAATTCAGAGATGATGATGAAAGAAAGGCTATGAGCTTGT 876  
QY 998 GGAATTCAGATATACAAACCTTCGATGATATCTCTATGTTACTATGTTGCTGGA 1057  
DB 877 GAAATTCAGATATACAAACCTTCGATGATATCTCTATGTTACTATGTTGCTGGA 936  
QY 1058 CTGTGATGATGATGATGTTCCAGTTATGAGTATGCACTGATCAAGGACACACAG 1117  
DB 937 CGGTGGGTTGATGATGATGTTCCAGTTATGAGTATGAGTATGAGTATGAGTATGAG 996  
QY 1118 AGAGTATATATATGCTGCTTTGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 1177  
DB 997 AGAGTATATATATGCTGCTTTGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 1056  
QY 1178 GAGATGATGAGAGATGCTCAGAGAGAGAGAGATATCTGCTCAAGATGATTTAGAC 1237  
DB 1057 GAGATGATGAGAGATGCTCAGAGAGAGAGAGATATCTGCTCAAGATGATTTAGAC 1116  
QY 1238 AGCAGGCTCTCCGACGAGACATATTTGCTGAGAGTGAATGATGATGAGAGAACT 1297  
DB 1117 AGCAGGCTCTATCCGATGAGATATATTTGCTGAGAGGCTGACGATTAATGAGAACT 1176  
QY 1298 TTAGAAGAAACAAATTCAGAGGCGAGAGAAATCTTGAATGATGATGAGAGAAAGTGTCA 1357  
DB 1177 TTAGAAGAAACAAATTCAGAGGCGAGAGAAATCTTGAATGATGATGAGAGAAAGTGTCA 1236  
QY 1358 CAGAACTGAGCTCTGCTAGTATGAGTGGCTGTGTTAACAGCCTGCTGTTGATGCAAGA 1417  
DB 1237 CAGAACTGAGCTCTGCTAGTATGAGTGGCTGTGTTAACAGCCTGCTGTTGATGCAAGA 1296  
QY 1418 TATTGACGAGATTGAGCCCAACGATCAACCACTTCAACAGAGAGGCTTATGTTAGCA 1477  
DB 1297 TACTATGATGATGAGTGAAGCCATGATCAACCACTTCAACAGAGAGATATGAGCA 1356  
QY 1478 AGCAGAG-AGGCTCTCACTGCGCATGCTGTTAGAGAAATCTGTTGCCCTTAAT 1536  
DB 1357 AATCAACCAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1416  
QY 1537 AGAATCTCTCTCCACTAGCAAGACATGAATGAAGTATGAGTCAATGATATATATAC 1596  
DB 1417 AATCTCTCTCTCTCAAGATATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1476  
QY 1597 ACT-----AAGAACTCAAGGATCTTGTAAATGAGATATCTTTGCTTAATGAT 1648  
DB 1477 AATGATCATTAAG 1536

QY 1649 CATCAAAAGTAGATTGT-----AAATTCAATATGACAACTCTGTGTAGAAATATTTCTCC 1704  
 Db 1537 CATAAATATATGTTGTAGTACATTCATTAATATCTCTTGTAGTGTGTATCTTC 1596  
 QY 1705 ACACACTCAATCAACCTCAAGTGAG 1728  
 Db 1597 ACTTATCTCAACTCTTTGAGAG 1620

## RESULT 5

US-08-300-582-2  
 ; Sequence 2, Application US/08300582  
 ; Patent No. 5750865  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BIRD, COLIN R.  
 ; APPLICANT: GRIERSON, DONALD  
 ; APPLICANT: SCHUCH, WOLFGANG W.  
 ; TITLE OF INVENTION: DNA, CONSTRUCTS, CELLS AND PLANTS  
 ; TITLE OF INVENTION: DERIVED THEREFROM  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: CUSHMAN DARBY & CUSHMAN  
 ; STREET: 1100 NEW YORK AVENUE, N.W.  
 ; CITY: WASHINGTON  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20005-3918  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/300,582  
 ; FILING DATE:  
 ; CLASSIFICATION: 800  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/859,523  
 ; FILING DATE: 12-AUG-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: KOKULIS, PAUL N.  
 ; REGISTRATION NUMBER: 16,773  
 ; REFERENCE/DOCKET NUMBER: 3893/95874  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 861-3000  
 ; TELEFAX: (202) 822-0944  
 ; TELEX: 6714627 CUSH  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1646 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; US-08-300-582-2

Query Match 48.7%; Score 842; DB 1; Length 1646;  
 Best Local Similarity 74.3%; Pred. No. 4.8e-232; Indels 92; Gaps 9;

Matches 1222; Conservative 0; Mismatches 330; Indels 92; Gaps 9;  
 QY 101 TATATCGTTGAATAATTAGTAGATAGCTCTAGTGATATCTAC--AAGTATGGTTT 157  
 Db 53 TAAATTTTGTGAGAGTGGAATATCTCTAGTGGAATCTAGAGATATTTATTT 112  
 QY 158 TGATTAATATAGGCTGAGTGAGAGTAACTAAAGAAACAAAACTTGGGATTT 217  
 Db 113 CTATTAACCTAAGTAAAGTTTGGAGGTGACA-AAAAGAAAGCAAAAATCTTGGATTT 171  
 QY 218 TTTGACACACGAGGTTCTTGTTCATGAGAGTCTGTGTTGTGTGGGTTT 277  
 Db 172 TTTGACACACCAAGGT--TTTCTGTCTCAGAGTGTGTGCTTGTATGGGTTT 228  
 QY 278 CTCACACTTCGAGGCTCGAATGGAGAGATGTTGTGATTCAGTCCGAGAAGAAAC 337

Db 229 CTC---CTTGACAGTCTCAAAATGGGACAGTTTCATGAAATCAGTCCGGAGGAAACC 285  
 QY 338 GCGCTTTTATATCATCCAGGTTCTCTAGTCTGAGATTTGAAATTTGATGGAAATGGGAA 397  
 Db 286 GTTTTATGATTCATCGAG-----CATAGAAATTTGGTGTCCAAATGAGGAA 333  
 QY 398 TCAAGAAAGGTGGAGACAAAGGTGAAATTTGGCTCTTTAATTTGATCCAAATATT 457  
 Db 334 TCAATAGAGT----- 344  
 QY 458 CATGCTGGGTGATCAAGACTGAAAGGAAAGCACTTCTCTGACAGTCCAGTTGG 517  
 Db 345 -----GGTGAAGACCAACTAATATAGACGGAATTTCTGTACGCTCTGCTATT 396  
 QY 518 TGGCTAGCCAGGCTGGAGAAATGACGTGTCTATCAGAGAAAAAGGTATGATGTGTAT 577  
 Db 397 TGGCTACTCATCTGGAGAACGGACATGACATCGAACAGATGTCTATGATGTGTAT 456  
 QY 578 TAAAGCAGCAGCTTTAGTGAAGAGCAGCTGAGATCTACCGATGATTTAGAAAGTGAAGC 637  
 Db 457 TGAGCAGCAGCAGCTTGTGTGAAGAGCAGCTGAGATCTACATGATTTAGAAAGTGAAGC 516  
 QY 638 CGATATTTGTTGTTCCAGGAAATTTGGGCTTTGAGTGAAGCATATGATGTTGGCG 697  
 Db 517 CGATATACCTAATTCGGGGAATTTGGGCTTTGAGTGAAGCATATGATGAGTGTG 576  
 QY 698 AAGTATGAGAGATGATGCAAGACATTTACTTAGGAAACCAAGCTAATACCCAGAGA 757  
 Db 577 AAGTATGAGAGATGATGCAAGACATTTACTTAGGAAACCTAATGATATATCCAGAGA 636  
 QY 758 GAAGAAAGCTATCTGGGCAATATATGTGTGTGTCAGAGAAAGCATGAGCTTTGTATG 817  
 Db 637 GAAGAAAGCTATCTGGGCAATATATGTGTGTGTCAGAGAAAGCATGAGCTTTGTATG 696  
 QY 818 GCCCTAATGATCCACATATCTCCGACGCTTTAGTATGATGGAGACAGGCTGGAAG 877  
 Db 697 GCCCAAGCATATATATATACCCGGCAGCTTGAATGATGGAGAAATAGGCTGAGAG 756  
 QY 878 ATATTTCAATGGGGCGGCTTTGATATGATGATGATGCTTTATCCGATATGCTGCCA 937  
 Db 757 ATGTTTCAATGGGGCGGCTTTGATATGATGATGATGCTTTGATCCGATATGCTTTCTA 816  
 QY 938 GATTTCTGTTGATATTCAGCCATTCAGAGATATGATTTGAAGAAATCGTATGCACTTGT 997  
 Db 817 ACTTTCAGTTGATATTCAGCCATTCAGAGATATGATTTGAAGAAATCGTATGCACTTGT 876  
 QY 998 GAAATCCAGATCAAAAATTTGATGAGCTATATCTTATTTGATATGTTGCTGTGA 1057  
 Db 877 GAAATCCAGATCAAAAATTTGATGAGCTATATCTTATTTGATATGTTGCTGTGA 936  
 QY 1058 CTGTAGATTTGATGAGGTGCTCCAGTTATGGTATTTGACCTGAAATCAAAAGCAACAG 1117  
 Db 937 CGTTGGGTTGATGAGGTGCTCCAGTTATGGTATTTGACCTGAAATCAAAAGCAACAG 996  
 QY 1118 AGAGTATATTAATGCTGTTGGCTTTAGGCTTGCATTAATCACTAATCAATATATCTCA 1177  
 Db 997 AGAGCTATATTAATGCTGTTGGCTTTAGGCTTGCATTAATCACTAATCAATATATCTCA 1056  
 QY 1057 GAGATGTTGAGAAAGTCCAGAAAGAGAGAGAGTCTTACCTCAAGATGAATTTAGCAC 1116  
 Db 1238 AGGAGGCTCTCCGACGAAGACATATTTCTGGAAGAGTGAAGTGAATGAGAGAACT 1297  
 QY 1117 AGGAGGCTCTATCCGAGGAAGATATATTTCTGGAAGAGTGAAGTGAATGAGAACT 1176  
 Db 1298 TTAATGAAGAAACAAATTCAGAGGCGAGAAATTTCTTGAATGAGTCAAGAAAGTGTCA 1357  
 QY 1177 TTATGAAGAAACAAATTCATATGAGGCAAGAAAGTCTTTGATGAGGCAAGAAAGCGTGA 1236  
 Db 1358 CAGAACTGACCTGTGATGATGAGTGGCTGTGTTAACAGGCTGCTGTTATGCAAGA 1417

Db 1237 CAGAAATGAGCTCAGCTAGTAAATTCCTGTATGGGCATCTTGTGTTGTAACCGCAAAA 1296  
Qy 1418 TATTGGACGAGATTGAAGCCCAAGCTACAACTTCCAGAGAGGGCTTATGTTAGCA 1477  
Db 1297 TACTAGATGAGATTGAAGCCCAAGCTACAACTTCCAGAGAGCATATGAGAGCA 1356  
Qy 1478 AGCCAAAG-AAGCTTCTCACTTGGCCATGCTTATGCAAAATCTCTTGGCCCCCTAT 1536  
Db 1357 AATCAAGAGAGTGTATGATTCATTACCTATGATGCAAAATCTCTTGGCCCCCTACA 1416  
Qy 1537 AGAATCTCTCTCCAGTACCAAGATGATGAGTGTAGTCAATGAGTATTTATG 1596  
Db 1417 AATGCGCTCTCTTCAAAAGATGAAGATATATATATATATATATATATATAC 1476  
Qy 1597 ACT-----AAAGAACTCAGTACTGTAAATGAGATCTTGTGTAATGCTAT 1648  
Db 1477 AATGATCATTAAGAAAAAGAAAGAAATGTTGTATGATATATATATATATAT 1536  
Qy 1649 CATCAAAAGTAAATGT-----AAATCAATATGACATCTTGTGTAATATTTTCTCC 1704  
Db 1537 CATTAATATTAAGTGTAGTAAATCATTAATATATATCTTGTATGTTGTATCTTC 1596  
Qy 1705 ACATCATTAACCTCAAGTAG 1728  
Db 1597 ACTTATCTCAACTCTTGTAGAG 1620

## RESULT 6

US-09-180-342-2  
; Sequence 2, Application US/09180342A  
; Patent No. 6239331  
; GENERAL INFORMATION:  
; APPLICANT: Drake, Caroline R.  
; APPLICANT: Bird, Colin R.  
; APPLICANT: Schuch, Wolfgang W.  
; TITLE OF INVENTION: Enhancement of Gene Expression  
; FILE REFERENCE: SEB0156  
; CURRENT APPLICATION NUMBER: US/09/180,342A  
; CURRENT FILING DATE: 1998-11-05  
; EARLIER APPLICATION NUMBER: PCT/GB97/01414  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: UK 9611981.3  
; EARLIER FILING DATE: 1996-06-07  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1239  
; TYPE: DNA  
; ORGANISM: Lycopersicon esculentum  
US-09-180-342-2

Query Match 46.9%; Score 811.2; DB 3; Length 1239;  
Best Local Similarity 83.8%; Pred. No. 3e-223;  
Matches 917; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

Qy 466 GGTGATCAAGAACTGAAAGGAAGCACTTCTGTACAGTCCAGTTGAGGCTAGC 525  
Db 145 GGTGAAAGCAACTATATATAGACGGAATTTTCTGTAAGGCTGTCTATTTTGGCTACT 204  
Qy 526 CCAGCTGAGAAATGACTGTGTATCATAGAGAAAAAGTGTATGATGTGTATTAAGCAG 585  
Db 205 CCATCTGAGAGACGAGCATGATCGAAGACAGATGTCTATGATGTGTATTTGAGGCG 264  
Qy 586 GAGCTTTAGTAAAGGACGCTGAGATTTACCGATGATTGAAGTGAAGCCGATATT 645  
Db 265 GAGGCTTGTGTAAAGGACGCTGAGATTTACCGATGATTGAAGTGAAGCCGATATTA 324  
Qy 646 GTTGTCCAGGGAATTTGGGCTTGTGATGAGCAATATGATGTTGTCGGAAGTAT 705  
Db 325 CCTATTTCCGGGGAATTTGGGCTTGTGATGAGCAATATGATGTTGTCGGAAGTAT 384  
Qy 706 GAGAGTATGAAAGCAATTTTACTTAGAACCAAGCTAATGACCCGAGAGAAAGAGA 765

Db 385 GCAGATATGCAAGACCTTTAACTTAGAACTATGCTAATGATCTCCGAGAGAAAGG 444  
Qy 766 GCTATCTGGGCAATATATATGTTGGGAGAGAAACGATAGAGCTTGTATGAGCCCTAT 825  
Db 445 GCTATCTGGGCAATATATATGTTGGGAGAGAAACGATAGAGCTTGTATGAGCCCTAAC 504  
Qy 826 GCATCCCAATTAACCTCCGCAAGCTTTAGATAGTGGAGACAGCCGCTGAAGATATTTTC 885  
Db 505 GCATCATATATTAACCCCGAGCCTTATAGTGGGAAAATAGCTGAAGATGTTTTC 564  
Qy 886 AGTGGCGGCAATTTGATATGCTGTGATGCTGCTTTATCCGATCTGTCTCAAGTTTCT 945  
Db 565 AATGGCGGCAATTTGATATGCTGTGATGCTGCTTTATCCGATCTGTCTCAAGTTTCTCA 624  
Qy 946 GTTGTATTCAGCAATTAAGATATGATGTAAGAAACGATAGCACTTGTGAAATCC 1005  
Db 625 GTTGTATTCAGCAATTAAGATATGATGTAAGAAACGATAGCACTTGTGAAATCC 684  
Qy 1006 AGATCAAAACTTTCGATGAGCTATATCTTATGTTTACTATGTTGCTGTACTGTAGGA 1065  
Db 685 AGATCAAAACTTTCGATGAGCTATATCTTATGTTTACTATGTTGCTGTACTGTAGGA 744  
Qy 1066 TTGATGAGTGTCCAGTTATGAGTATGTCACCTGATCAAGCAAGCAAGAGAGTGA 1125  
Db 745 TTGATGAGTGTCCAGTTATGAGTATGTCACCTGATCAAGCAAGCAAGAGAGTGA 804  
Qy 1126 TATATGCTGCTTGTGGCTTTAGGGCTTGCATATCACTTAACCATATATCTAGAGATGA 1185  
Db 805 TATATGCTGCTTGTGGCTTTAGGGCTTGCATATCACTTAACCATATATCTAGAGATGA 864  
Qy 1186 GGAGAGATGTCAGAGAGAGAGAGTATCTTGCCTCAAGATGAAATGAGCAGGAGAGG 1245  
Db 865 GGAGAGATGTCAGAGAGAGAGAGTATCTTGCCTCAAGATGAAATGAGCAGGAGAGT 924  
Qy 1246 CTCTCCGAGAGACATATTTCTGGAAGAGTATGTAAGTGAAGTGAAGCACTTTATGAAG 1305  
Db 925 CTATCCGATGAAGATATATTTGCTGGAAGGCTGACCGATATATGAGATCTTTATGAAG 984  
Qy 1306 AAACAAATTCAGAGGCGAGAGAAATCTTGTATGATGAGAGAAAGTGTACAGAACTG 1365  
Db 985 AAACAAATTCAGAGGCGAGAGAAATCTTGTATGAGGAGAGAAAGGCTACAGAAATG 1044  
Qy 1366 GACTCTGCTAGTATGAGGCTGTGTAAACAGCGCTGTGTGTATGCAAGATATTTGAG 1425  
Db 1045 AGCTGAGTATGATATTTCCCTGTATGAGCACTTTGTGCTTGTACCGAAATATCTAGAT 1104  
Qy 1426 GAGATGAGGCAAGCACTTACCAACACTTCAAGAGAGGCTTATGTTAGCAAGCCAAAG 1485  
Db 1105 GAGATGAGGCAAGCACTTACCAACACTTCAAGAGAGGCTTATGTTAGCAAGCCAAAG 1164  
Qy 1486 AAGCTTTCACCTTGGCCATTTGCTTATGCAAAATCTTGTGCCCCCTAATGAACTTCC 1545  
Db 1165 AAGTGTATGATTAACCTATGCAATGCAAAATCTTGTGCCCCCTAATGAACTTCC 1224  
Qy 1546 TCTCCACTAGCAAA 1559  
Db 1225 TCTCTCAAAAGATA 1238

## RESULT 7

US-08-579-667-3  
; Sequence 3, Application US/08579667  
; Patent No. 5703624  
; GENERAL INFORMATION:  
; APPLICANT: Fitzmaurice, Wayne P.  
; APPLICANT: Hellmann, Gary M.  
; APPLICANT: Grill, Laurence K.  
; APPLICANT: Kumagai, Monto H.  
; APPLICANT: Della-Cioppa, Guy R.  
; TITLE OF INVENTION: DNA SEQUENCES ENCODING ENZYMES USEFUL IN  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:

```

ADDRESS: Virginia C. Bennett
STREET: 1211 East Morehead Street, PO Drawer 34009
CITY: Charlotte
STATE: No. 5705624th Carolina
COUNTRY: USA
ZIP: 28234

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/579, 667
APPLICATION NUMBER: US/08/579, 667
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Virginia C.
REGISTRATION NUMBER: 37,092
REFERENCE/DOCKET NUMBER: 627-196
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1316 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1239
US-08-579-667-3

Query Match      46.9%; Score 811.2; DB 1; Length 1316;
Best Local Similarity 80.2%; Pred. No. 3e-223;
Matches 1039; Conservative 0; Mismatches 173; Indels 84; Gaps 4;

QY 250  ATGCTGTGCTTGTGTTGGGTTGTTCTCCCACTTCGAGGTCTCGAATGGGACAGCA 309
DB 1  ATGCTGTGCTTGTGTTGGGTTGTTCTCCCACTTCGAGGTCTCGAATGGGACAGCA 57
QY 310  TTGTTGATTCAGTCCGAGAGAAACCGCGTCTTTGTATCATCCAGGTTCTTACCTGA 369
DB 58  TTCTTGATTCATCCGGAGAGAAACCGGTTTGTATGGTCCAGG-----105
QY 370  GATGAGATTTGATGTGAAATGGGAATCAAGAAAGTGGGACAAAGTGAATTTT 429
DB 106  CATGGAATTTAGTGTGCAATGAGAGAAACAAGAGGTGTGGAACAATGTGGAAATTTT 165
QY 430  GGCCTTTAATTCGTGATCCAAGATTCATGCTTGGGTGATCAAGAACTGAAAGGGA 489
DB 166  GGTTCTGTAA-----175
QY 490  AGCACTTCTCTGTACAGTCCAGTTGGTGGCTACGCCAGCTGAGAAAT---GACTGTG 546
DB 176  -----AGTCTGCTATGTGTGCTACACCGGGGGGAGAAATGGCCAGCATG 219
QY 547  TCATAGAGAAAAGGTGTATGTGTGTTAAAGCAGGCACTTTAGTGAAGGCGAG 606
DB 220  ACATCAGAACAGATGTTATGATGTGTGTTAAACAACAGCTTTAGTGAAGGCGAG 279
QY 607  CTGAGATCTACCGATGATTTGAAGTGAAGCGGATATTGTTGTTCCAGGGGAATTTGGGC 666
DB 280  TTGAGATCTACGATGATTTGAAGTGAAGCGGAGATCCCTCTCCCGGGGAATTTGAGC 339
QY 667  TTGTTAGTGAAGCATGATCGTTGTGCGCAAGATGTGCAAGATGACAAAGCATTT 726
DB 340  TTGTTAGTGAAGCATGATGATGTGTGCGCAAGATGTGCAAGATGACAAAGCATTT 399
QY 727  TACTTGAAGAACCACTATATGACCCCAAGAGAGAAAGCTATCTGGGCAATATATGTG 786
DB 400  TACTTGAAGAACCTATATGACTCCAGAGAGAAAGGCTATTTGGGCAATATATATTA 459

```

```

QY 787  TGTGACAGAGAACGATGAGCTGTGTGATGGCTTAATGCATCCCATTAATCCGCA 846
DB 460  TGTGACAGAGAACGATGAGCTGTGTGATGGCTTAATGCATCCCATTAATCCGCA 519
QY 847  GCTTTAGATAGTGGGAGACCGAGCTGGAGATATTTTCACTGGGCGCCATTGATATG 906
DB 520  GCCTTAGATAGTGGGAGAACCGAGCTGGAGATATTTTCACTGGGCGCCATTGATATG 579
QY 907  CTGATCTGCTTATCCGATGATGCTCCAGATTCCTGTATGATTTACCATTCAGCA 966
DB 580  CTGATCTGCTTATCCGATGATGCTCCAGATTCCTGTATGATTTACCATTCAGCA 639
QY 967  GATATGATGAGAGATCCGATGAGCTTGTGAAATCCAGATACAAACCTTTGATGAG 1026
DB 640  GATATGATGAGAGATCCGATGAGCTTGTGAAATCCAGATACAAACCTTTGATGAG 699
QY 1027  CTATATCTATTTGTTATCTATGTTGCTGTGATCTGTGAGATTTGATGATGCTTCA 1086
DB 700  CTATATCTATTTGTTATCTATGTTGCTGTGATCTGTGAGATTTGATGATGCTTCA 759
QY 1087  GGTATGCACTGATGATGAGAGCAAGAGAGTGTATATATGCTGCTTGGCTTTA 1146
DB 760  GGTATGCACTGATGATGAGAGCAAGAGAGTGTATATATGCTGCTTGGCTTTG 819
QY 1147  GGGCTTCAATCACTAATCAATATATCTCAGAGATGTAGAGAGATGCGAGAGAGA 1206
DB 820  GGTATGCGAATCACTAATCAATATCTCAGAGATGTAGAGAGATGCGAGAGAGA 879
QY 1207  AGAGATATCTTGTCTCAAGATGATTTGCAAGAGGAGGCTCTCCGAGAGACATTTT 1266
DB 880  AGAGATATCTTGTCTCAAGATGATTTGCAAGAGGAGGCTCTCCGAGAGACATTTT 939
QY 1267  GCTGGAAGATGATGATGAGAGAGCTTTATGAAAGAAACAATTCAGAGGCGAGG 1326
DB 940  ACTGGAAGATGATGATGAGAGAGCTTTATGAAAGAAACAATTCAGAGGCGAGA 999
QY 1327  AAATTTCTTGTATGATGATGAGAGAGAGTGTCAAGAACTGAGCTCTGTAGATGAG 1386
DB 1000  AAGTTCTTCAATGAGAGAGAGAGAGTGTCAAGAACTGAGCTCTGTAGATGAG 1059
QY 1387  GTGTTAAGAGGCTGCTGTGTGTATCCGAGATTTGAGAGAGATGAGCAAGCATAC 1446
DB 1060  GTATGGGCACTTGTGCTGTGTATCCGCAATATCTCAGCAGATGCAAGCAATGACTAC 1119
QY 1447  AACCACTTCAAGAGAGGCTTATGTTAGCAAGCAAGAAAGCTTCACCTTGGCCATT 1506
DB 1120  AACCACTTCAAGAGAGGCTTATGTTAGCAAGCAAGAAAGCTTATTCCTTACCTATT 1179
QY 1507  GCTTATGCAAAATCTCTTGTGCCCCCTAATGAACT 1542
DB 1180  GCTTATGCAAAATCTCTTGTGCCCCCTAATGAACT 1215

RESULT 8
US-09-180-342-1
Sequence 1, Application US/09180342A
Patent No. 6239331
GENERAL INFORMATION:
APPLICANT: Drake, Caroline R.
APPLICANT: Bird, Colin R.
APPLICANT: Schuch, Wolfgang W.
TITLE OF INVENTION: Enhancement of Gene Expression
FILE REFERENCE: SRE50156
CURRENT APPLICATION NUMBER: US/09/180,342A
EARLIER FILING DATE: 1998-11-05
EARLIER APPLICATION NUMBER: PCT/GB97/01414
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: UK 9611981.3
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1

```



LENGTH: 1239  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA  
 US-09-180-342-1

Query Match 24.0%; Score 415.2; DB 3; Length 1239;  
 Best Local Similarity 61.6%; Pred. No. 1.9e-109;  
 Matches 663; Conservative 0; Mismatches 413; Indels 0; Gaps 0;

466 GGTGATCAAGAACTGAAAAGGAGACCTTTCTCTGATCAAGTCCAGTTTGGCTAGC 525  
 145 GGAGGTAAACAGACAACACGATGAAAGTTCTCAGTTAGATCAACAATCCTTGACACA 204  
 526 CCAGCTGAGAAATGCTGTGATCAGAGAAAAGGTATGATGATGATTAATAAGCAG 585  
 205 CTTAGCGGTGAGAGAACTATGATGACGAGCAAAATGATGATGATGATGATGATGAT 264  
 586 GCAGCTTTAGTGAAGAGGAGCTGATGATGATGATGATGATGATGATGATGATGAT 645  
 265 GCTGACTATTAAGTAAAGTATGATGATGATGATGATGATGATGATGATGATGAT 324  
 646 GTTGTCCAGGAAATTTGGCTTTGTTGATGAAGCATATGATGATGATGATGATGAT 705  
 325 CCAAATACCTGGAACCTTGACTCTTCTTGAGGCTTACAGACAGATGCGAGAGGTTGC 384  
 706 GCAGAGTATGCAAGACATTTTACTAGAACCAAGCTAATGACCCAGAGAGAAAGAGA 765  
 385 GGAGATATGCTAAACCTTCAATTTGGTACCATGTTGATATACACCAAAAGGCTGCT 444  
 766 GCTATCTGGGCAATATATGTTGTCGACAGAGACGATGATGATGATGATGATGAT 825  
 445 GCAATATGGGCTATTTAGCTTTGTTGAGCGGATGATGATGATGATGATGATGAT 504  
 826 GCATCCACATATCTCCGCAAGCTTTAGATGATGATGATGATGATGATGATGATGAT 885  
 505 GCTAGTTACATTAACCCCTGCTCTGACAGATGAGAGAACGTTTGGAGACGCTTT 564  
 886 AGTGGGCGCCATTTGATATGCTTGAATGCTTATCCGATACCTGCCAGATTTCT 945  
 565 AACGGAGACCTTTGATATGTTGACGAGACCTTATGATGATGATGATGATGATGAT 624  
 946 GTTGTATTTACGCCATTCAGATATGATGATGATGATGATGATGATGATGATGAT 1005  
 625 GTGACATCCATCTTTGCGGACATGATGACAGGCGATGATGATGATGATGATGAT 684  
 1006 AGATACAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065  
 685 CATTATTAAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 744  
 1066 TTGATGAGTGTCCAGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1125  
 745 CTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 804  
 1126 TATTAATGCTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1185  
 805 TACACCGACGACCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 864  
 1186 GGAGAGATGCGAGAGAGAGAGATATATCTGCTCAAGATGATGATGATGATGATGAT 1245  
 865 GGTGAGAGACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 924  
 1246 CTCTCCGAGAGAGACATATTTCTGAGAGAGTGAATGATGATGATGATGATGATGAT 1305  
 925 TTGATGATGAGAGACATTTTCCAGATGATGATGATGATGATGATGATGATGATGAT 984  
 1306 AAAACAATTCAGAGGCGAGAGAAATCTTGAATGATGATGATGATGATGATGATGATGAT 1365  
 985 AAGCAGATTCACCGTCTGTTAAATTTTTCAGAGAGCTGAAAGAGAGATTCATGACT 1044  
 1366 GACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1425

Db 1045 TCTAGTGCATCAAGGTTTCCAGTTTGGCCAGCCCTTGCTCTATAGAAAGATTTTGAC 1104  
 Oy 1426 GAGATTGAAGCAACAGATCAACAACTTCAAGAGAGGCTTATGATGAGCAAG 1485  
 Db 1105 GAATCGAGGCTAACGATTAATTAATTTTCTAAGACGCTTACGTTTCTTAAGAGCAA 1164  
 Oy 1486 AAGCTTCTACCTTGCCCATGCTTATGCAAAATCTCTTGCCCCCTAATAGAAC 1541  
 Db 1165 AACTTATGCTGCTTCCAAATCGCTTAAGAGCTTGTTCCACCAACTAAGAC 1220

# RESULT 9 US-08-579-667-9

Sequence 9, Application US/08579667  
 Patent No. 5705624

GENERAL INFORMATION:  
 APPLICANT: Fitzmaurice, Wayne P.

APPLICANT: Hellmann, Gary M.

APPLICANT: Grill, Laurence K.

APPLICANT: Kumagai, Monto H.

APPLICANT: Della-Cioppa, Guy R.

TITLE OF INVENTION: DNA SEQUENCES ENCODING ENZYMES USEFUL IN

TITLE OF INVENTION: PHYTOENE BIOSYNTHESIS

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESS: Virginia C. Bennett

STREET: 1211 East Morehead Street, PO Drawer 34009

CITY: Charlotte

STATE: No. 5705624th Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/579,667

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Bennett, Virginia C.

REGISTRATION NUMBER: 37,092

REFERENCE/DOCKET NUMBER: 627-196

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-420-2200

TELEFAX: 919-881-3175

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 749 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

## US-08-579-667-9

Query Match 10.0%; Score 172; DB 1; Length 749;  
 Best Local Similarity 65.7%; Pred. No. 1.2e-39;  
 Matches 393; Conservative 0; Mismatches 110; Indels 95; Gaps 6;

109 TTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 168  
 226 TAGACTCTAGTGGGAATCTATGAGATGATTTTATTTTATTTTATTTTATTTTAA 295  
 Oy 169 GCTGAGTGAAGGTAACATTA-----GGAAAGACAAAACCTTGGAAATTTT 220  
 Db 296 GCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 355  
 Oy 221 AGACCAAGGAGGATTTCTTTTTCATGAGCATGCTGTTGTTGTTGTTGTTGTTGTT 280  
 Db 356 AGACAAACCAAGGATTTTGTGTTTC--AGAAATGCTGTTGCTTGTATGAGGTTGTTAC 412  
 Oy 281 CCATTCGAGGCTCGAATGGAAGAGATTTGTTGATTCAGTCCGAGAGGAAACCGCG 340

```

Db      413  ---CTTGAGGCTCAATGGGACAGATTCTTGATTCAGGAAGGAGGAAACCGGG 469
Qy      341  TCTTGTATCAATCCAGGTTCTTCACTCGAATAGAAATTTGATGTGGAATGGAAATCA 400
Db      470  TTTTGTACTCGTCGGG-----CATGGAATTTAGTGTGCAATAGAGAAATCA 517
Qy      401  AGAAGGTGGAGCAAAAGTGAATTTTGCTTTTATTCGTATCAAGATATTCAT 460
Db      518  AAAAGGCTGTAACAAAGTGAATTTTG----- 548
Qy      461  GCTTGCTGATCAAGAACTGAAAAGGAGACATTTCTCTGTACAGTTCAGTTGGTG 520
Db      549  -----TTCCGTACGCTCGATGCTGG 571
Qy      521  CTAGCCAGCTGAGAAAT---GACTGTCTCATCAGAAAAAGGTGATATGCTGAT 577
Db      572  CTACACCAACGGGAAATGCGCAATGACATCGAAACAGAAAGTTTATGATGTGAT 631
Qy      578  TAAAGCAGGACCTTGAAGAGCAGCTGAGATCTACCGATGATTTAGAAGTGAAGC 637
Db      632  TGAACCAAGCAGCTTGAAGAGGAGCAGCTGAGATCTACGTATGATTTAGAAGTGAAGC 691
Qy      638  CGGATATTTGTTTCCAGGGAATTTGGCTTGTGAGTGAAGCATATGATGCTTGG 695
Db      692  CGGATATCTCTCTCCCGGGAATTTGAGCTTGTAAAGTGAAGCATATGATGCTGATG 749

```

RESULT 10  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.

APPLICANT: SCHEFLINGER, F.

TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/232,463

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/935,313

FILING DATE:

APPLICATION NUMBER: EP 91 114 300.6

FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 30472/114 IMMUNO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300

TELEFAX: (703)836-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:  
; CLONE: pTZgpc-Fls  
; US-08-232-463-14

Query Match 3.1%; Score 54; DB 1; Length 7218;  
Best Local Similarity 3.1%; Pred. No. 2.8e-05;  
Matches 12; Conservative 223; Mismatches 153; Indels 0; Gaps 0;

```

Qy      1093  GCACCTGAATCAAGGACAAACAGAGGTATATATCTGCTTGGCTTTAGGCTT 1152
Db      1445  GAATTTGCTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1386
Qy      1153  GCAATCAATCAATCAATATATCTAGAGATGTAGAGAGATGCCAGAGAGAGTA 1212
Db      1385  RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1326
Qy      1213  TACTTGCTCAAGATGATTAAGCAGGCGGCTCTCCAGCAAGACATATTTGCTGA 1272
Db      1325  RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1266
Qy      1273  AGAGTACTGATTAAGTGAAGGAACTTATGAGAAACAATTCAGAGGAGGAAATTC 1332
Db      1265  RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1206
Qy      1333  TTGATGAGTCAGAGAAAGGTGTCAAGAACTGACTCTGATGATGAGCTGTGTTA 1392
Db      1205  RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1146
Qy      1393  ACAGCGCTGCTGTTGATTCGACAGATTTGACAGAGATTGAAGCAACGACTACAC 1452
Db      1145  RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1086
Qy      1453  TTCACAGAGAGGCTTATGTTGCAAGC 1480
Db      1085  RRRRRRRRRRRRRRRRRRRATCGAAGC 1058

```

RESULT 11  
US-08-260-546-10

; Sequence 10, Application US/08260546

; Patent No. 5922602

GENERAL INFORMATION:

APPLICANT: Kumagai, Monte H.

APPLICANT: della-Cioppa, Guy R.

APPLICANT: Donson, Jonathan

APPLICANT: Harvey, Damon A.

TITLE OF INVENTION: The Cytoplasmic Inhibition of Gene

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 2710 Sand Hill Road

CITY: Menlo Park

STATE: California

COUNTRY: U.S.A.

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/260,546

FILING DATE: 16-JUN-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Halluin, Albert P.

REGISTRATION NUMBER: 25,227

REFERENCE/DOCKET NUMBER: 8129-086-999

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-854-3660

TELEFAX: 415-854-3694

TELEX: 66141 PENNIR

```

; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..114
US-08-260-546-10

Query Match
Best Local Similarity 79.8%; Score 52.4; DB 2; Length 114;
Matches 75; Conservative 0; Mismatches 16; Indels 3; Gaps 1;

QY 245 TGAGATGTCGTCCTTGTGTTGGGTTGTTCCACCTTCGAGGTTCTGATGGGA 304
DB 23 TAAATATGTCGTGTCCTGTGTTATGCGTTGTTCTC---CTTGTGACGCTTCAAAATGGGA 79

QY 305 CAGATTGTGATTCAGTCCGAGGAAGAAACCG 338
DB 80 CAAATTTCATGGAATCAGTCCGAGGAGGAACCG 113

RESULT 12
US-09-436-068A-12
; Sequence 12, Application US/09436068A
; Patent No. 6376752
; GENERAL INFORMATION:
; APPLICANT: KUMAGAI, MONTO H.
; APPLICANT: della-CIOPPA, GUY R.
; APPLICANT: DONSON, JONATHAN
; APPLICANT: HARVEY, DAMON A.
; APPLICANT: GRILL, LAURENCE K.
; TITLE OF INVENTION: THE CYTOPLASMIC INHIBITION OF GENE
; FILE REFERENCE: 008010086US02
; CURRENT APPLICATION NUMBER: US/09/436,068A
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 09/265,576
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 08/260,546
; PRIOR FILING DATE: 1994-06-16
; PRIOR APPLICATION NUMBER: 08/184,237
; PRIOR FILING DATE: 1994-01-19
; PRIOR APPLICATION NUMBER: 07/923,692
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 114
; TYPE: DNA
; ORGANISM: Tomato mosaic virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128) ... (114)
US-09-436-068A-12

Query Match
Best Local Similarity 79.8%; Score 52.4; DB 4; Length 114;
Matches 75; Conservative 0; Mismatches 16; Indels 3; Gaps 1;

QY 245 TGAGATGTCGTCCTTGTGTTGGGTTGTTCCACCTTCGAGGTTCTGATGGGA 304
DB 23 TAAATATGTCGTGTCCTGTGTTATGCGTTGTTCTC---CTTGTGACGCTTCAAAATGGGA 79

QY 305 CAGATTGTGATTCAGTCCGAGGAAGAAACCG 338
DB 80 CAAATTTCATGGAATCAGTCCGAGGAGGAACCG 113

RESULT 13
```

```

US-09-265-576-12
; Sequence 12, Application US/09265576
; Patent No. 6479291
; GENERAL INFORMATION:
; APPLICANT: KUMAGAI, MONTO H.
; APPLICANT: della-CIOPPA, GUY R.
; APPLICANT: DONSON, JONATHAN
; APPLICANT: HARVEY, DAMON A.
; APPLICANT: GRILL, LAURENCE K.
; TITLE OF INVENTION: THE CYTOPLASMIC INHIBITION OF GENE
; FILE REFERENCE: 008010086US02
; CURRENT APPLICATION NUMBER: US/09/265,576
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 09/436,068
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 09/265,576
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 08/260,546
; PRIOR FILING DATE: 1994-06-16
; PRIOR APPLICATION NUMBER: 08/184,237
; PRIOR FILING DATE: 1994-01-19
; PRIOR APPLICATION NUMBER: 07/923,692
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 114
; TYPE: DNA
; ORGANISM: Tomato mosaic virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128) ... (114)
US-09-265-576-12

Query Match
Best Local Similarity 79.8%; Score 52.4; DB 4; Length 114;
Matches 75; Conservative 0; Mismatches 16; Indels 3; Gaps 1;

QY 245 TGAGATGTCGTCCTTGTGTTGGGTTGTTCCACCTTCGAGGTTCTGATGGGA 304
DB 23 TAAATATGTCGTGTCCTGTGTTATGCGTTGTTCTC---CTTGTGACGCTTCAAAATGGGA 79

QY 305 CAGATTGTGATTCAGTCCGAGGAAGAAACCG 338
DB 80 CAAATTTCATGGAATCAGTCCGAGGAGGAACCG 113

RESULT 14
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 16, 2004, 03:01:56 ; Search time 1327 Seconds  
(without alignments)  
4589.895 Million cell updates/sec

Title: US-09-847-081B-1

Perfect score: 1728  
Sequence: 1 agaaaccagaagaagacaac.....tcatcaaccctcaagttag 1728

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1728	100.0	1728	US-09-847-081B-1	Sequence 1, Appli
2	857.8	49.6	1712	US-09-847-081B-3	Sequence 3, Appli
3	806.4	46.7	1239	US-09-371-307-75	Sequence 75, Appli
4	806.4	46.7	1239	US-10-401-321-75	Sequence 75, Appli
5	578.4	33.5	1269	US-09-938-842A-729	Sequence 729, App
6	177.4	10.3	684	US-10-149-759-53	Sequence 53, Appli
7	52.4	3.0	114	US-10-236-508-9	Sequence 9, Appli
8	52.4	3.0	114	US-10-103-450-12	Sequence 12, Appli
9	52.4	3.0	114	US-10-103-450-12	Sequence 12, Appli
10	52.4	3.0	114	US-10-137-765-9	Sequence 9, Appli
11	52.4	3.0	114	US-10-146-337-9	Sequence 9, Appli
12	42.8	2.5	948	US-10-156-761-1642	Sequence 1642, Ap
13	42.8	2.5	9025608	US-10-156-761-1	Sequence 1, Appli
14	41.6	2.4	3673778	US-10-312-841-2	Sequence 2, Appli
15	41.2	2.4	534	US-09-736-457-1310	Sequence 1310, Ap

c 16	41.2	2.4	534	10	US-09-902-941-1310	Sequence 1310, Ap
c 17	41.2	2.4	534	10	US-09-849-626-1310	Sequence 1310, Ap
c 18	41.2	2.4	534	13	US-10-113-872-1310	Sequence 1310, Ap
c 19	41.2	2.4	534	15	US-10-017-754-1310	Sequence 1310, Ap
c 20	41.2	2.4	400	10	US-09-938-842A-2919	Sequence 2919, Ap
c 21	40.8	2.4	6087	13	US-10-311-455-391	Sequence 391, App
c 22	40.4	2.3	7657	13	US-10-311-455-1995	Sequence 1995, App
c 23	40.4	2.3	7657	15	US-10-239-676-185	Sequence 185, App
c 24	40.2	2.3	6378	15	US-10-242-056-58	Sequence 58, Appli
c 25	40.2	2.3	3016	13	US-10-311-455-1217	Sequence 1217, Ap
c 26	39	2.3	281	10	US-09-969-373-264	Sequence 264, App
c 27	39	2.3	513509	11	US-09-754-853A-4	Sequence 4, Appli
c 28	38.8	2.2	671	15	US-10-184-634-246	Sequence 346, App
c 29	38.8	2.2	1716	9	US-09-815-242-6522	Sequence 6522, Ap
c 30	38.6	2.2	16173	13	US-10-240-485-215	Sequence 21, Appli
c 31	38.4	2.2	6000	13	US-10-311-455-1105	Sequence 1105, Ap
c 32	38.2	2.2	6195	13	US-10-311-455-563	Sequence 563, App
c 33	38.2	2.2	142299	11	US-09-911-077A-14	Sequence 14, Appli
c 34	38.2	2.2	7040	15	US-10-172-086-13	Sequence 13, Appli
c 35	38	2.2	15416	13	US-10-311-455-2204	Sequence 2204, Ap
c 36	37.8	2.2	867	13	US-10-140-472-20	Sequence 20, Appli
c 37	37.8	2.2	867	13	US-10-141-761-20	Sequence 20, Appli
c 38	37.8	2.2	867	13	US-10-142-885-20	Sequence 20, Appli
c 39	37.8	2.2	867	13	US-10-158-790-20	Sequence 20, Appli
c 40	37.8	2.2	867	13	US-10-137-871-20	Sequence 20, Appli
c 41	37.8	2.2	867	13	US-10-140-805-20	Sequence 20, Appli
c 42	37.8	2.2	867	13	US-10-140-864-20	Sequence 20, Appli
c 43	37.8	2.2	867	13	US-10-140-923-20	Sequence 20, Appli
c 44	37.8	2.2	867	13	US-10-141-756-20	Sequence 20, Appli
c 45	37.8	2.2	867	13	US-10-141-756-20	Sequence 20, Appli

## ALIGNMENTS

RESULT 1  
US-09-847-081B-1  
; Sequence 1, Application US/09847081B  
; Patent No. US2002012846A1  
; GENERAL INFORMATION:  
; APPLICANT: BAYER AG  
; TITLE OF INVENTION: DNA encoding the tobacco phytoene synthase  
; FILE REFERENCE: Le A 34 326  
; CURRENT APPLICATION NUMBER: US/09/847,081B  
; CURRENT FILING DATE: 2001-05-02  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1728  
; TYPE: DNA  
; ORGANISM: Nicotiana tabacum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (244)..(1566)  
US-09-847-081B-1

Query Match	100.0%	Score 1728;	DB: 10;	Length 1728;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1728;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	AGAAACCCAGAAAGAACAGAGTTTCTTCTTGATGAGTCATTTGGCTGCTT	60	
DB	1	AGAAACCCAGAAAGAACAGAGTTTCTTCTTGATGAGTCATTTGGCTGCTT	60	
QY	61	GTGTAAGCCAAAGTCGTTCACTTTCTTATATCCGATTTTATATCGTTGAATTAAG	120	
DB	61	GTGTAAGCCAAAGTCGTTCACTTTCTTATATCCGATTTTATATCGTTGAATTAAG	120	
QY	121	GATTAACCTCTGATGATATCTTCAAGATTTGTTTGAATAAATAGCGTAGAGTGGA	180	
DB	121	GATTAACCTCTGATGATATCTTCAAGATTTGTTTGAATAAATAGCGTAGAGTGGA	180	
QY	181	AGGTAACTAAGAAAGAAAGAAAGAAAGCTGGAGATTTGTTTGAACACCGAGTTCTTGT	240	

```

Db      181 AGGTAACTAAAGGAAACAAAACCTGGGAATGTTTAGACCAACGAGATCTTGT 240
Qy      241 TTCAATGACATGCTGCTGCTTGTGTGGGTGTTTCCCACTCCGAGTCTCGAAT 300
Db      241 TTCAATGACATGCTGCTGCTTGTGTGGGTGTTTCCCACTCCGAGTCTCGAAT 300
Qy      301 GGGACAGATGTTGGATTCAATCCAGAGAAACCCGCTTTGTATCAATCCAGTTC 360
Db      301 GGGACAGATGTTGGATTCAATCCAGAGAAACCCGCTTTGTATCAATCCAGTTC 360
Qy      361 CTAGCTCGAGATAGGAATTTGATGGAATGGAGAAATCAAGAAAGTGGAGACAAAG 420
Db      361 CTAGCTCGAGATAGGAATTTGATGGAATGGAGAAATCAAGAAAGTGGAGACAAAG 420
Qy      421 TGGAAATTTGGCTCTTTAATTTGCTGATCCAGATATTTATCTGGGGGATCAAGAACT 480
Db      421 TGGAAATTTGGCTCTTTAATTTGCTGATCCAGATATTTATCTGGGGGATCAAGAACT 480
Qy      481 GAAAAGGAGACATTTCTCTGTACAGTCCAGTTGGTGGCTAGGCCAGCTGAGAAATG 540
Db      481 GAAAAGGAGACATTTCTCTGTACAGTCCAGTTGGTGGCTAGGCCAGCTGAGAAATG 540
Qy      541 ACTGTGTCATCAGAGAAAAAGGTGTATGATGTTAAAGCAGGACGCTTTAGTGAAG 600
Db      541 ACTGTGTCATCAGAGAAAAAGGTGTATGATGTTAAAGCAGGACGCTTTAGTGAAG 600
Qy      601 AGGACGTCGAGATCTACCGATGATTTAGAAAGTGAAGCCGATATTTGTTCCAGGGAAT 660
Db      601 AGGACGTCGAGATCTACCGATGATTTAGAAAGTGAAGCCGATATTTGTTCCAGGGAAT 660
Qy      661 TTGGGCTTGTGAGTGAAGCATATGATGTTGGCGAAGATGTCAGAGATATGCAAAAG 720
Db      661 TTGGGCTTGTGAGTGAAGCATATGATGTTGGCGAAGATGTCAGAGATATGCAAAAG 720
Qy      721 ACATTTTACTTGAAGCAACGATATGACCCGAGAGAGAAAGATATCTGCGCAATA 780
Db      721 ACATTTTACTTGAAGCAACGATATGACCCGAGAGAGAAAGATATCTGCGCAATA 780
Qy      781 TATGTGTGTGTCAGAGAAACGATGAGTGTGTGAATGTCCTTAATGCAATCAATACT 840
Db      781 TATGTGTGTGTCAGAGAAACGATGAGTGTGTGAATGTCCTTAATGCAATCAATACT 840
Qy      841 CCGCAAGCTTATAGATAGGAGAGACAGGCTGGAAGATATTTCAAGTGGGCGCATTT 900
Db      841 CCGCAAGCTTATAGATAGGAGAGACAGGCTGGAAGATATTTCAAGTGGGCGCATTT 900
Qy      901 GATATGCTTGAATGCTGCTTTATCCGATATCTGTCAGATTTCTGTTGATATTCAGCA 960
Db      901 GATATGCTTGAATGCTGCTTTATCCGATATCTGTCAGATTTCTGTTGATATTCAGCA 960
Qy      961 TTCAGAGATATGATTTGAAGGAATGCTATGAGCTTTGTGGAATCCAGATATCAAACTTC 1020
Db      961 TTCAGAGATATGATTTGAAGGAATGCTATGAGCTTTGTGGAATCCAGATATCAAACTTC 1020
Qy      1021 GATGAGCATATCTCTATTTGTTATCTATGTTGCTGTAATGATGATGATGATGATGATGAT 1080
Db      1021 GATGAGCATATCTCTATTTGTTATCTATGTTGCTGTAATGATGATGATGATGATGATGAT 1080
Qy      1081 GTTATGAGTATTTGCACTGATCAAAAGCAACAGAGATGATATATATGCTGCTTGG 1140
Db      1081 GTTATGAGTATTTGCACTGATCAAAAGCAACAGAGATGATATATATGCTGCTTGG 1140
Qy      1141 GCTTTAGGGCTTGCATCAATCAATCAATATCTCAGAGATGTAAGGAAGATGCCAGA 1200
Db      1141 GCTTTAGGGCTTGCATCAATCAATCAATATCTCAGAGATGTAAGGAAGATGCCAGA 1200
Qy      1201 AGAGAGAGATATCTTGGCTCAAGATGATATTTAGACAGGAGGCTCCGAGAGAGAC 1260
Db      1201 AGAGAGAGATATCTTGGCTCAAGATGATATTTAGACAGGAGGCTCCGAGAGAGAC 1260
Qy      1261 ATATTTGTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320

```

```

Db      1261 ATATTTGCTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Qy      1321 GCGAGAAATTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
Db      1321 GCGAGAAATTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
Qy      1381 TGGGCTGTGTTAACAGGCTGCTGTGTGTATCGCAAGATATTTGAACGATGTAAGCCAAC 1440
Db      1381 TGGGCTGTGTTAACAGGCTGCTGTGTGTATCGCAAGATATTTGAACGATGTAAGCCAAC 1440
Qy      1441 GACTACAACTTCAACAGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db      1441 GACTACAACTTCAACAGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Qy      1501 CCCATGCTTATGCAAAATCTCTTGTGCCCCCTTAATGAACTTCTCTCACTAGCAAG 1560
Db      1501 CCCATGCTTATGCAAAATCTCTTGTGCCCCCTTAATGAACTTCTCTCACTAGCAAG 1560
Qy      1561 ACATGATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Db      1561 ACATGATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Qy      1621 AATGATATCTTTTCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db      1621 AATGATATCTTTTCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Qy      1681 ATCTCTGTGTAATATTTCTCCACATCAACCTCAAGTGAAG 1728
Db      1681 ATCTCTGTGTAATATTTCTCCACATCAACCTCAAGTGAAG 1728

RESULT 2
US-09-847-081B-3
: Sequence 3, Application US/09847081B
: Patent No. US20020128464A1
: GENERAL INFORMATION:
: APPLICANT: BAYER AG
: TITLE OF INVENTION: DNA encoding the tobacco phytoene synthase
: FILE REFERENCE: Le A 34 326
: CURRENT APPLICATION NUMBER: US/09/847, 081B
: CURRENT FILING DATE: 2001-05-02
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO 3
: LENGTH: 1712
: TYPE: DNA
: ORGANISM: Nicotiana tabacum
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (333)..(1565)
: FEATURE:
: NAME/KEY: unsure
: LOCATION: 135, 139
: OTHER INFORMATION: Xaa is unknown or other
: NAME/KEY: unsure
: LOCATION: 51
: OTHER INFORMATION: n can be any nucleotide
US-09-847-081B-3

Query Match      49.6%; Score 857.8; DB 10; Length 1712;
Best Local Similarity 76.7%; Pred. No. 9, 8e-234;
Matches 1206; Conservative 2; Mismatches 265; Indels 100; Gaps 9;

```



QY	706	GCAGAGTATGCAGAAACATTTTACTTAGAGAACCAAGCTAATGACCCCAAGAGAAACA	765
Db	385	GCAGAGTATGCAGAAACGTTTAACTTAGAACCTATGCTAATGACTCCAGAGAGAAAG	444
QY	766	GCTATCTGGGCATATATATGTGTGTGCAGAGAACGATGACCTGTGTATGGCCCTAAT	825
Db	445	GCTATCTGGGCATATATATGTATGTGTGCAGAAACAGATGAACCTGTGTATGGCCCAAC	504
QY	826	GCATCCCACTAATCTCCGACCTTTAGTATAGTGGGAAACCAAGGTGGAAATATTTTC	885
Db	505	GCATATATATTTATCCCGGACCCCTAGATAGGTGGAAATATGGCTAAGATGTTTTTC	564
QY	886	AATGGGCGGCCCATTTTGATATGCTTGATGCTGCTTTATCCGATACTGCTCCAGATTTCC	945
Db	565	AATGGGCGGCCCATTTTGATGCTGCTTGATGCTGCTTTGTCGATATCAATTTCTTCC	624
QY	946	GTTGATATTTCCACCCATTCAGAAATATGATTTGAAAGAAATGCGTATGGAATCTTGAGAAATCC	100
Db	625	GTTGATATTTCCACCCATTCAGAAATATGATTTGAAAGAAATGCGTATGGAATCTTGAGAAATG	684
QY	1006	AGATACAAAACCTTTGATGAGCTATATCTATTTGTAATGATGTTCTGTAAGTA	106
Db	685	AGATACAAAACCTTTGATGAGCTATATCTATTTGTAATGATGTTCTGTAAGCTTTGG	744
QY	1066	TTGATGAGTGTTCAGATTATGGGTAATGCACTGAATCAAGAGCAACAAGAGTGT	112
Db	745	TTGATGAGTGTTCAGATTATGGGTAATGCGCCCTGATCAAGAGCAACAAGAGAGCT	804
QY	1126	TATATATGCGCTTTAGGCTTTAGGGCTGGCAATTCATTAACCAATTTACTTCAGAGATGT	118
Db	805	TATATATGCTGCTTTAGGCTTTAGGGATCGCAATTCATTAACCTAATCTCAGAGATGT	864
QY	1186	GGAGAAATGCCAGAAAGAAAGATATCTTGCTCAAGATGAATTTAGCAAGCAGGG	124
Db	865	GGAGAAATGCCAGAAAGAAAGATCTACTTGCTCAAGATGAATTTAGCAAGCAGGG	924
QY	1246	CTCTCCGACGAAGATATTTGCTGGAAGAGTACGTATTAATGGAGAACTTATTAAG	130
Db	925	CTATCCGATGAAGATATATTTGCTGGAAGGGTGACCGATTAATGGAGAACTTATTAAG	984
QY	1306	AAACAAATTCAGAGGCGAGAGAAATCTTTGATGATGATCAGAGAAAGTGTCAAGAACTG	136
Db	985	AAACAAATTCAGAGGCGAGAGAAATCTTTGATGATGATGAGAAATGGCGTGACAGAAATG	104
QY	1366	GACTCTGCTAGTATGATGAGCTGTGTTAACAGCGCTCTGTGTATGCGAAATATTGGAC	142
Db	1045	AGCTCAGCTAGTATATTCCTGTATATGGGATCTTTGTGTTGTAACGCAAAATATCTAGAT	110
QY	1426	GAGATTTGAAGCCAGACTACAAACAATTCACAAAGAGGGCTTATTTAGCAAGCCAAAG	148
Db	1105	GAGATTTGAAGCCAGACTACAAACAATTCACAAAGAGAGCATATGTGAGCAAAATCAAAG	116
QY	1486	AAGCTTCTCACTTGCCCATTTGCTTATGCAAAATCTCTTGTCGCCCTAATAGAACTTCC	154
Db	1165	AAGTTGATTTGCAATTCCTATTTGCAATAGCAAAATCTCTTGTCGCCCTCAAAAACTGCC	122
QY	1546	TCTCCACTAGCAAA 1559	
Db	1225	TCTCTTCAAGATTA 1238	

```

1  APPLICANT: Ream, Joel E.
2  APPLICANT: Logusch, Sherry J.
3  APPLICANT: Baerson, Scott R.
4  TITLE OF INVENTION: Methods for Controlling Gibberellin Levels
5  FILE REFERENCE: 11899.0216.DVUS01 (MOBT:216--1)
6  CURRENT APPLICATION NUMBER: US/10/401,321
7  CURRENT FILING DATE: 2003-03-27
8  NUMBER OF SEQ ID NOS: 89
9  SOFTWARE: PatentIn version 3.2
10 SEQ ID NO 75
11 LENGTH: 1239
12 TYPE: DNA
13 ORGANISM: Lycopersicon esculentum
14 US-10-401-321-75

```

Query Match	46.7%	Score 806.4	DB 12	Length 1239
Best Local Similarity	83.5%	Pred. No. 3,76-219		
Matches	914	Conservative	0	Mismatches 180; Indels 0; Gaps 0
QY	466	GGTGGATCAAGAACTGAAAAAGGGAAGCACTTCTCTGTGACGTCCAGTTGGTGTGAC	525	
Db	145	GGTGAAGAGCAACTAATTAATGACGGAAAAATTTTCTGTAACGCTGTGCTAATTTGGCTACT	204	
QY	526	CCAGCTGGAGAAATGACTGTGTCTATCAAGAAAAAGGTGTTATGATGTGTATTAAGCAG	585	
Db	205	CCAACTTGGAGAACGAGAGATGACATCCGAACAGATGGCTATGATGTGTGTTTGAGGGCAG	264	
QY	586	GCAGCTTTAGGAAGAGGAGCTAGATCTCCCATGATTTTGAAGAGTGAACCCGGATATT	645	
Db	265	GCAGCTTTGGTGAAGAGGCACTGAGATCTACCAATGATTTGAAGTGAACCCGGATATTA	324	
QY	646	GTTGTTCCAGGAAATTTTGGGCTTGTGTAAGTGAACATATGATCGTTGTGCGAAGTATGT	705	
Db	325	CCATATCCGGGAATTTTGGGCTTGTGTAAGTGAACATATGATGATGTGTGTAAGTATGT	384	
QY	706	GCAGAGTATCGAAAGCACTTTTACTTTAGAACCCAGCTAAAGACCCGAGAGAAAGAA	765	
Db	385	GCAGAGTATCGAAAGAGGTTTAACTTTAGGAACATATGCTAAAGACTCCCGAAGAAAGAG	444	
QY	766	GCTATCTGGGCAATATATGTGTGTGAGAGAAACGATGAGCTTGTGTATGAGCCCTAAT	825	
Db	445	GCTATCTGGGCAATATATGTATGTGTGAGAAAGAACATGATGAACCTGTGTATGAGCCCAAC	504	
QY	826	GCATCCCAATTAATCTCGCAGCTTTAGATATGTGGAGACCAAGGCTGGAAATATTTTC	885	
Db	505	GCATCAATATATTACCCCGGAGCCCTTAGATAGTGGGAAAAATAGGCTTAGAAGATGTTTTTC	564	
QY	886	AGTGGGGGGCATTGTAATGCTTGAATGCTGTATTCGATCTACTGTCTCCAGATTTCCT	945	
Db	565	AATGGGGGGCCATTGTAATGCTGTGAATGCTGTATTCGATCTACTGTCTCCAGATTTCCTCA	624	
QY	946	GTTGATATTGAGCATTACAGAGATATATTTGAAGAAATGCGTATGSACTTTGSAATCC	1005	
Db	625	GTTGATATTGAGCATTACAGAGATATATTTGAAGAAATGCGTATGSACTTTGSAAAAAATCG	684	
QY	1006	AGATACAAAACTTTGATGAGCTATATCTATTGTTATCTATGTTGCTGTGTACTGTAGGA	1065	
Db	685	AGATACAAAAAACTTCGACGAACCTATACCTTATTGTTATTATGTTGCTGTGTAACGGTTGGG	744	
QY	1066	TTGATGAGCTTTCCAGTTATGGGTAATGCAACCTGAATCAAGGGAACAACAGAGAGCTTA	1125	
Db	745	TTGATGAGCTTTCCAAATTAATGGGTAATGCGCCCTGAATCAAGGGAACAACAGAGAGCTTA	804	
QY	1126	TATTAATCTGCTTTGAGCTTTAGGGCTTGCAAATCAACTAACCAATATATCTCAGAGATGTA	1185	
Db	805	TATTAATCTGCTTTGGCTCTGGGATGCGAAATCAATTAATCAATATCTCAGAGATGTTT	864	
QY	1186	GGAGAGATGTCGAGAGAGAGAAAGATATCTTGTCTCAAGATGATTTGCAACAGGACGGG	1245	
Db	865	GGAGAGATGTCGAGAGAGAGAAAGATGCTAATCTTGCTCAAGATGATTTGCAACAGGACAGT	924	
QY	1246	CTCTCCGACGAAGACATATTTGCTGGAAAGATGATCTAATAGTGGAGAACTTTATGAG	1305	



Db 925 CTATCCGATGAAGATATATTTGCTGAGAGGCTGACCGATTAATGAGAACTTTTATGAAG 984  
 Qy 1306 AACAATAATTGAGAGGCGAGAAATCTTGTAGTACAGAGAAAGGTGTCAAGACTG 1365  
 Db 985 AACAATAATGATAGGCAAGAAAGTTCTTGTAGTACAGAGAAAGGTGTCAAGACTG 1044  
 Qy 1366 GACTCTGCTAGTATGAGTGGCTGTGTATACAGCGCTGTGTGTATGCAAGATATTGAC 1425  
 Db 1045 AGCTCAGCTAGTATATTCCTGTATGCGCATCTTTGGTCTTGTATCCGCAAAATATAGAT 1104  
 Qy 1426 GAGATTGAGCCAGACATCAACAACTTCAAGAGAGGCTTATGTCAGAACCAAG 1485  
 Db 1105 GAGATTGAGCCAGATCAACAACTTCAAGAGAGCAATATGTCAGAACCAAG 1164  
 Qy 1486 AAGCTTCTCAGCTTGGCCATGCTTATGCAAAATCTTGTGCCCCCTATATGAACCTTC 1545  
 Db 1165 AAGTTGATTCATTAATGCTTATGCAAAATCTTGTGCCCCCTATATGAACCTTC 1224  
 Qy 1546 TCTCCACTAGCAAA 1559  
 Db 1225 TCTCTTCAAGATA 1238

## RESULT 5

US-09-938-842A-729  
 ; Sequence 729, Application US/09938842A  
 ; Patent No. US20020160378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Krepes, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 ; FILE REFERENCE: SCRIPT300-3  
 ; CURRENT APPLICATION NUMBER: US/09/938,842A  
 ; PRIOR FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/227,866  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/264,647  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ; PRIOR FILING DATE: 2001-06-22  
 ; NUMBER OF SEQ ID NOS: 5379  
 ; SEQ ID NO 729  
 ; LENGTH: 1269  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 ; US-09-938-842A-729

Query Match 33.5%; Score 578.4; DB 10; Length 1269;  
 Best Local Similarity 72.5%; Pred. No. 4.7e-154;  
 Matches 766; Conservative 0; Mismatches 281; Indels 9; Gaps 1;

Qy 474 AAGAACTGAAAAGGAGCACTTCTCTGTACAGTCCAGTTGGTGGTAGCCAGCTGG 533  
 Db 177 AAGGAACCGAAGTAGAGAAATGGTGTGTCTTCAAGCTTAGTAGCAAGTCTTCTGG 236  
 Qy 534 AGAATGACTGTGTCTATCAGAGAAAAGGTATGATGTGTATTAAGCAGGCACTTT 593  
 Db 237 AGAGATAGCTCTTTTCTTGAAGAGAGTTTACATGTTGTGTAACAAGCTGCTTT 296  
 Qy 594 AGTGAAGGAGAGCTGAGATCTAC-----CGATGATTTGAAGTGAACCGCATAT 644  
 Db 297 GGTGAACCAACAGCTTAGGCTCTTCTTATGACCTTGATGTAAGAAAACAAGATGT 356  
 Qy 645 TGTGTTCCAGGAAATTTGGGCTTGTGAGTGAAGATATGATCGTTGGGGAAGTATG 704  
 Db 357 TGTTCCTCTCGGAGATTGATGTTGTGTTGGGTGAAGCTTATGATCGATGGGGAAGTTTG 416  
 Qy 705 TGCAGATATGCAAGACATTTTACTTAGAAACCAAGCTATGAGCCCAAGAGAAAG 764  
 Db 417 CGCTGAATATGCTAAGACGTTTATCTTGAACTTGTCTTATGACACCCGAAGGCGA 476

Qy 765 AGCTATCTGGCAATATATGTGTGTGAGAGAGAAAGATGAGTGTGTGATGGCCCTAA 824  
 Db 477 GCGCATTTGGGCAATCTACGTTTGTGTGTAAGAACTGATGAACCTTGTGTGATGGCCCA 536  
 Qy 825 TGCATCCACATTAATCCGCAAGCTTTAGATAGTGGAGAACAGCTGGAAGATATTTT 884  
 Db 537 TGTCTACATATTAATCCCAATGCTTATGATGATGAGTGAAGCAAGATTAAGATCTTTT 596  
 Qy 885 CAGTGGGGGCGCATTTGATATGCTTGTATGCTCTTATCCGATATCTGTCTCCAGATTTCC 944  
 Db 597 CGGTGTGCTCTTGTGATATGCTTGTATGCTGTCTCTGCTGATACAGTTGATATCC 656  
 Qy 945 TGTGATATATGAGCATTCAAGATATGATATGAAAGAAATGCTATGCACTTGTGAAATC 1004  
 Db 657 GGTGATATATGAGCATTCAAGATATGATATGAAAGAAATGATGATGATGATGATGAT 716  
 Qy 1005 CAGATACAAAATTTGATGAGCTATATCTTATTTTACTATATGTTGCTGTACTGTAG 1064  
 Db 717 GAGATACCAAAATTTGATGATCTATATCTTACTATGATGATGATGATGATGATGAT 776  
 Qy 1065 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1124  
 Db 777 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 836  
 Qy 1125 ATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1184  
 Db 837 TTAACAGCTGCTTGTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 896  
 Qy 1185 AGGAGAGATGCGAGAGAGAGAGATATCTTGTGCTGATGATGATGATGATGATGATGAT 1244  
 Db 897 AGGAGAGATGCGAGAGAGAGAGATATCTTGTGCTGATGATGATGATGATGATGATGAT 956  
 Qy 1245 GCTTCCGAGAGAGAGATATCTTGTGCTGATGATGATGATGATGATGATGATGATGAT 1304  
 Db 957 TCTTCAATGAGAGAGAGAGATATCTTGTGCTGATGATGATGATGATGATGATGATGAT 1016  
 Qy 1305 GAAACAAATTCAGAGAGAGAGAGATATCTTGTGATGATGATGATGATGATGATGATGAT 1364  
 Db 1017 AATGAGCTTAAAGAGAGAGAGATATCTTGTGATGATGATGATGATGATGATGATGAT 1076  
 Qy 1365 GCACTCTCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1424  
 Db 1077 CAGTCCGCTAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1136  
 Qy 1425 CGAGATGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1484  
 Db 1137 CGAGATGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1196  
 Qy 1485 GAACTTCTCACTTGGCCATTTGCTTATGCAAAATC 1520  
 Db 1197 GAAATTCAGCTTTTGGCATTTGCTTATGCTTAATC 1232

## RESULT 6

US-10-149-759-53  
 ; Sequence 53, Application US/10149759  
 ; Publication No. US20030157592A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lersch, Jens  
 ; APPLICANT: Renz, Andreas  
 ; APPLICANT: Ehrhardt, Thomas  
 ; APPLICANT: Reindl, Andreas  
 ; APPLICANT: Clippus, Petra  
 ; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins  
 ; TITLE OF INVENTION: Involved in the synthesis of tocopherols and  
 ; FILE REFERENCE: BASF/NAE 1333/99 PCT/US  
 ; CURRENT APPLICATION NUMBER: US/10/149,759  
 ; PRIOR FILING DATE: 2002-10-17  
 ; PRIOR APPLICATION NUMBER: PCT/EP/00/12698  
 ; NUMBER OF SEQ ID NOS: 82

SOFTWARE: WordPerfect version 6.1  
 SEQ ID NO 53  
 LENGTH: 684  
 TYPE: DNA  
 ORGANISM: Physcomitrella patens  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (2)..(397)  
 OTHER INFORMATION: 02\_pprocl\_046\_a07rev  
 US-10-149-759-53

Query Match 10.3%; Score 177.4; DB 13; Length 684;  
 Best Local Similarity 68.0%; Pred. No. 9.5e-40;  
 Matches 247; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 1171 ATACTGAGATGTAGAGAGATGCCAGAGAGAGATATCTTCCCTCAAGTAA 1230  
 DB 5 ATCCAGAGAGATGTAGAGAGATGACGCCGTGACAGATATCTCCACAGATGAA 64  
 QY 1231 TTACACAGAGAGAGCTCTCCGACGACATATTTGTGGAAGAGTACTGATAGTG 1290  
 DB 65 CTGGACAGCTTCGGTCTCTCGATGACACATTTTGTGGAAGATGATGATTAAG 124  
 QY 1291 AGGAATTTATGAGAAACAATTCAGAGGCGAGAAATTTCTTGATGACTCAGAGAA 1350  
 DB 125 AGGCATTCATGAGAAAGACCAATTAAGAGCTAGAGTCTTTGTGAGAGCTGAGAAA 184  
 QY 1351 GGTGTACAGAGAGCTGCTGTAGTATGAGGCTGTGTAAAGCGCTGCTGTGTAT 1410  
 DB 185 GGTGTACGTGAGCTGAGAAAGACGTGCTGCTGTGTGTGCTGCTTCAATCTTTAC 244  
 QY 1411 CGCAGATATTTGACGAGATTTGAAGCAAGCACTTCAACAAGAGAGGCTTAT 1470  
 DB 245 CAGGAATTCGTGAGAGCCATTTGAAGCAAGCACTTCAACAAGAGAGCTTAC 304  
 QY 1471 GTTAGCAAGCAAGAGCTTCTACCTTGGCCATTTGCTTAATGCAAAATCTTGTGCCC 1530  
 DB 305 GTAGCAAGAGTGAAGAAAGCTGCTTCTACCTTAATGCGCTTAAGCAGAGCGTGTGCTCA 364  
 QY 1531 CCT 1533  
 DB 365 CCT 367

RESULT 7  
 US-10-236-508-9  
 Sequence 9, Application US/10236508  
 Publication No. US20030167512A1  
 GENERAL INFORMATION:  
 APPLICANT: Monto H. Kumagai  
 APPLICANT: Guy R. della-Cioppa  
 APPLICANT: Robert L. Erwin  
 APPLICANT: David R. McGee  
 TITLE OF INVENTION: METHOD OF DETERMINING THE PRESENCE OF A TRAIT IN A PLANT BY  
 TITLE OF INVENTION: TRANSFECTING A NUCLEIC ACID SEQUENCE OF A DONOR PLANT INTO A  
 FILE REFERENCE: 008010137US05  
 CURRENT APPLICATION NUMBER: US/10/236,508  
 CURRENT FILING DATE: 2002-09-06  
 NUMBER OF SEQ ID NOS: 48  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 9  
 LENGTH: 114  
 TYPE: DNA  
 ORGANISM: Tomato mosaic virus  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (28)...(114)  
 US-10-236-508-9

Query Match 3.0%; Score 52.4; DB 13; Length 114;  
 Best Local Similarity 79.8%; Pred. No. 0.00015;  
 Matches 75; Conservative 0; Mismatches 16; Indels 3; Gaps 1;

QY 245 TGACATGCTGTGCTTTGTGTTGAGTTTCTCCACTTCGAGCTTCGAATGGA 304  
 DB 23 TAATATGCTGTGCTTGTCTGTATGAGGTGTTCTC---CTTGACGTCTCAATGGA 79  
 QY 305 CAGATTGTTGATTCAGTCCGAGAGAAACCG 338  
 DB 80 CAAGTTTCATGATCAGTCCGGAGGAAACCG 113

RESULT 8  
 US-10-103-450-12  
 Sequence 12, Application US/10103450  
 Publication No. US20030219897A9  
 GENERAL INFORMATION:  
 APPLICANT: KUMAGAI, MONTO H.  
 APPLICANT: DELLA-CIOPPA, GUY R.  
 APPLICANT: DONSON, JONATHAN  
 APPLICANT: HARVEY, DAMON A.  
 APPLICANT: GRILL, LAURENCE K.  
 TITLE OF INVENTION: THE CYTOPLASMIC INHIBITION OF GENE  
 TITLE OF INVENTION: EXPRESSION BY VIRAL RNA  
 FILE REFERENCE: 008010086US02  
 CURRENT APPLICATION NUMBER: US/10/103,450  
 CURRENT FILING DATE: 2002-03-20  
 PRIOR APPLICATION NUMBER: 09/436,068  
 PRIOR FILING DATE: 1999-11-08  
 PRIOR APPLICATION NUMBER: 08/260,546  
 PRIOR FILING DATE: 1994-06-16  
 PRIOR APPLICATION NUMBER: 08/184,237  
 PRIOR FILING DATE: 1994-01-19  
 PRIOR APPLICATION NUMBER: 07/923,692  
 PRIOR FILING DATE: 1992-07-31  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 12  
 LENGTH: 114  
 TYPE: DNA  
 ORGANISM: Tomato mosaic virus  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (28)...(114)  
 US-10-103-450-12

Query Match 3.0%; Score 52.4; DB 13; Length 114;  
 Best Local Similarity 79.8%; Pred. No. 0.00015;  
 Matches 75; Conservative 0; Mismatches 16; Indels 3; Gaps 1;

QY 245 TGACATGCTGTGCTTTGTGTTGAGTTTCTCCACTTCGAGCTTCGAATGGA 304  
 DB 23 TAATATGCTGTGCTTGTCTGTATGAGGTGTTCTC---CTTGACGTCTCAATGGA 79  
 QY 305 CAGATTGTTGATTCAGTCCGAGAGAAACCG 338  
 DB 80 CAAGTTTCATGATCAGTCCGGAGGAAACCG 113

RESULT 9  
 US-10-103-450-12  
 Sequence 12, Application US/10103450  
 Publication No. US20020155605A1  
 GENERAL INFORMATION:  
 APPLICANT: KUMAGAI, MONTO H.  
 APPLICANT: DELLA-CIOPPA, GUY R.  
 APPLICANT: DONSON, JONATHAN  
 APPLICANT: HARVEY, DAMON A.  
 APPLICANT: GRILL, LAURENCE K.  
 TITLE OF INVENTION: THE CYTOPLASMIC INHIBITION OF GENE  
 TITLE OF INVENTION: EXPRESSION BY VIRAL RNA  
 FILE REFERENCE: 008010086US02  
 CURRENT APPLICATION NUMBER: US/10/103,450  
 CURRENT FILING DATE: 2002-03-20  
 PRIOR APPLICATION NUMBER: 09/436,068

PRIOR FILING DATE: 1999-11-08  
PRIOR APPLICATION NUMBER: 08/260,546  
PRIOR FILING DATE: 1994-06-16  
PRIOR APPLICATION NUMBER: 08/184,237  
PRIOR FILING DATE: 1994-01-19  
PRIOR APPLICATION NUMBER: 07/923,692  
PRIOR FILING DATE: 1992-07-31  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 12  
LENGTH: 114  
TYPE: DNA  
ORGANISM: Tomato mosaic virus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (28)...(114)  
US-10-103-450-12

Query Match  
Best Local Similarity 79.8%; Score 52.4; DB 14; Length 114;  
Matches 75; Conservative 0; Mismatches 16; Indels 3; Gaps 1;

QY 245 TGAGCATGCTCTGTTGCTTTGTTGGTGTGTTCTCCCACTTCCGAGGTCCTGAATGGGA 304  
DB 23 TAAATATGCTGCTGTTGCTTTATGAGGTGTTCTC---CTGTGACGTCCTCAATGGGA 79  
QY 305 CAGGATTTGTTGATTCAGTCCGAGGAAGAAACCG 338  
DB 80 CAAGTTTCATGGAATCAGTCCGAGGAAGAAACCG 113

RESULT 10  
US-10-137-765-9

Sequence 9, Application US/10137765  
Publication No. US20030028926A1  
GENERAL INFORMATION:  
APPLICANT: KUMAGAI, Monto H.  
APPLICANT: DELLA-CIOPPA, Guy R.  
APPLICANT: ERWIN, Robert L.  
APPLICANT: MCGEE, David R.  
TITLE OF INVENTION: METHOD OF COMPILING A FUNCTIONAL GENE PROFILE BY  
TITLE OF INVENTION: TRANSECTING A NUCLEIC  
TITLE OF INVENTION: ACID SEQUENCE OF A NON-PLANT DONOR INTO A HOST PLANT IN A POSITIVE  
FILE REFERENCE: 008010137US07  
CURRENT APPLICATION NUMBER: US/10/137,765  
CURRENT FILING DATE: 2002-05-01  
PRIOR APPLICATION NUMBER: US/09/359,300A  
PRIOR FILING DATE: 2002-04-29  
NUMBER OF SEQ ID NOS: 71  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 9  
LENGTH: 114  
TYPE: DNA  
ORGANISM: Tomato mosaic virus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (28)...(115)  
US-10-137-765-9

Query Match  
Best Local Similarity 79.8%; Score 52.4; DB 15; Length 114;  
Matches 75; Conservative 0; Mismatches 16; Indels 3; Gaps 1;

QY 245 TGAGCATGCTCTGTTGCTTTGTTGGTGTGTTCTCCCACTTCCGAGGTCCTGAATGGGA 304  
DB 23 TAAATATGCTGCTGTTGCTTTATGAGGTGTTCTC---CTGTGACGTCCTCAATGGGA 79  
QY 305 CAGGATTTGTTGATTCAGTCCGAGGAAGAAACCG 338  
DB 80 CAAGTTTCATGGAATCAGTCCGAGGAAGAAACCG 113

RESULT 11  
US-10-146-337-9

Sequence 9, Application US/10146337  
Publication No. US20030041355A1  
GENERAL INFORMATION:  
APPLICANT: KUMAGAI, Monto H.  
APPLICANT: DELLA-CIOPPA, Guy R.  
APPLICANT: ERWIN, Robert L.  
APPLICANT: MCGEE, David R.  
TITLE OF INVENTION: METHOD OF COMPILING A FUNCTIONAL GENE PROFILE BY  
TITLE OF INVENTION: TRANSECTING A NUCLEIC  
TITLE OF INVENTION: ACID SEQUENCE OF A NON-PLANT DONOR INTO A HOST PLANT IN A POSITIVE  
FILE REFERENCE: 008010137US07  
CURRENT APPLICATION NUMBER: US/10/146,337  
CURRENT FILING DATE: 2002-05-14  
PRIOR APPLICATION NUMBER: US/09/359,300  
PRIOR FILING DATE: 2002-04-29  
NUMBER OF SEQ ID NOS: 71  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 9  
LENGTH: 114  
TYPE: DNA  
ORGANISM: Tomato mosaic virus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (28)...(115)  
US-10-146-337-9

Query Match  
Best Local Similarity 79.8%; Score 52.4; DB 15; Length 114;  
Matches 75; Conservative 0; Mismatches 16; Indels 3; Gaps 1;

QY 245 TGAGCATGCTCTGTTGCTTTGTTGGTGTGTTCTCCCACTTCCGAGGTCCTGAATGGGA 304  
DB 23 TAAATATGCTGCTTCCCTGTTATGAGGTGTTCTC---CTGTGACGTCCTCAATGGGA 79  
QY 305 CAGGATTTGTTGATTCAGTCCGAGGAAGAAACCG 338  
DB 80 CAAGTTTCATGGAATCAGTCCGAGGAAGAAACCG 113

RESULT 12  
US-10-156-761-1642

Sequence 1642, Application US/10156761  
Publication No. US20030119018A1  
GENERAL INFORMATION:  
APPLICANT: OMURA, SATOSHI  
APPLICANT: IKEDA, HARUO  
APPLICANT: ISHIKAWA, JUN  
APPLICANT: HORIKAWA, HIROSHI  
APPLICANT: SHIBA, TADAYOSHI  
APPLICANT: SAKAKI, YOSHIYUKI  
APPLICANT: HATTORI, MASAHIRA  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-262  
CURRENT APPLICATION NUMBER: US/10/156,761  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: JP 2001-204089  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: JP 2001-272697  
PRIOR FILING DATE: 2001-08-02  
NUMBER OF SEQ ID NOS: 15109  
SEQ ID NO 1642  
LENGTH: 948  
TYPE: DNA  
ORGANISM: Streptomyces avermitilis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(948)  
US-10-156-761-1642

Query Match  
2.5%; Score 42.8; DB 15; Length 948;



; CURRENT FILING DATE: 2000-12-13  
 ; NUMBER OF SEQ ID NOS: 1864  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1310  
 ; LENGTH: 534  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)..(534)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-736-457-1310

Query Match 2.4%; Score 41.2; DB 10; Length 534;  
 Best Local Similarity 57.9%; Pred. No. 0.62;  
 Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy	97	TTTTTAATCGTTGAAATTAGTGATGACTCTAGTGTATCTACAAATTTGGTTT	156
Db	465	TATTTACAACTTTATCATAGTAAATGTAATAATTAATCTATGATTTT	406
Qy	157	TTGATAAATAGGCTGAGGTGAGGTACATAAAGCAAAACTTGGGAATTG	216
Db	406	ATACAAAATCTCTAAGTACAAATGTAAATGCTGAATAAAACATTTCACCTT	346
Qy	217	TTTTAG 222	
Db	345	TTCCAG 340	

Search completed: January 16, 2004, 06:30:02  
 Job time : 1348 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 00:54:54 ; Search time 3744 Seconds  
(without alignments)  
11217.451 Million cell updates/sec

Title: US-09-847-081b-1

Perfect score: 1728  
Sequence: 1 agaaaccagaagaacaac.....tcatcaaacctcaagtgg 1728

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estcov:\*  
6: em\_estcpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	655.4	37.9	781	13	B0511016 EST18431
2	621.6	36.0	751	12	B0046203 EST1595321
3	605.2	35.0	787	12	BM410846 EST1585173
4	568.4	32.9	677	10	BG351357 104A12 Ma

Result No.	Score	Query Match	Length	ID	Description
5	552.2	32.0	686	9	AM442101 EST311497
6	551.2	31.9	765	12	BM408984 EST583311
7	541.2	31.3	682	12	BM409200 EST583527
8	532.8	30.8	672	9	AM442407 EST311803
9	532.2	30.2	642	10	BE433198 EST399727
10	515.8	29.8	676	9	AM222027 EST298838
11	498.2	28.8	634	12	BM412533 EST586860
12	487.2	28.2	599	12	BI921390 EST441293
13	486.4	28.1	600	10	BE460889 EST412308
14	483.4	28.0	623	10	BE432511 EST399040
15	478.8	27.7	685	9	AM221932 EST28743
16	474.4	27.4	819	13	BU636463 EST300127
17	473.2	27.4	582	9	AM223316 EST300127
18	466.8	27.0	582	9	AM223666 EST300477
19	466.4	27.0	581	10	BE432955 EST399484
20	466.4	27.0	623	12	BM536249 EST589271
21	461.4	26.7	791	9	AM223528 EST300339
22	458.4	26.5	573	10	BE435308 EST406386
23	458.4	26.5	605	9	BE435064 EST406142
24	458.2	26.5	585	9	AM222245 EST299056
25	458	26.5	1655	11	AY11032 Zea mays
26	457.8	26.5	565	10	BE432595 EST399124
27	453.2	26.2	562	10	BE433966 EST405044
28	449.4	26.0	581	9	AM222806 EST299617
29	442.4	25.6	556	9	AM222152 EST28963
30	439.6	25.4	542	10	BE431550 EST36365
31	439	25.4	543	12	BM411875 EST586202
32	437.4	25.3	634	9	AM441216 EST310612
33	435	25.2	555	10	BE460463 EST411882
34	434.4	25.1	566	10	BF112979 EST440662
35	431.6	25.0	542	10	BE432748 EST399277
36	429.2	24.8	563	10	BE461396 EST412815
37	427.8	24.8	580	9	AM092680 EST285870
38	427.6	24.7	563	10	BE437095 EST408213
39	426.8	24.7	536	10	BE434911 EST405989
40	425.4	24.6	558	9	AM222427 EST29238
41	425.4	24.6	1201	11	AY108547 Zea mays
42	424.2	24.5	744	12	BM412719 EST587057
43	423.6	24.5	551	10	BE432084 EST398613
44	415.2	24.0	536	9	AM223238 EST300049
45	409.2	23.7	519	10	BE433537 EST400066

## ALIGNMENTS

RESULT 1  
B0511016/c  
LOCUS  
DEFINITION EST618431 Generation of a set of potato CDNA clones for microarray analyses mixed potato tissues Solanum tuberosum CDNA clone STM067 3' end, mRNA sequence.

ACCESSION B0511016  
VERSION B0511016.2 GI:21926690  
KEYWORDS  
SOURCE EST.  
ORGANISM Solanum tuberosum (potato)

REFERENCE Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C., Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and Karameycheva, S.A.  
1 (bases 1 to 781)

On Jun 10, 2002 this sequence version replaced gi:21369885.  
Unpublished  
Generation of a set of potato CDNA clones for microarray analyses

Other ESTs: EST618430  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
This clone can be obtained from the University of Arizona Genomics

Institute. Orders can be made through URL:  
<http://genome.arizona.edu/orders/>  
 Seq primer: T7.

FEATURES  
 SOURCE  
 location/Qualifiers  
 1..781

/organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec or Bintje"  
 /db\_xref="taxon:4113"  
 /clone="STM067"  
 /tissue\_type="mixed tissues"  
 /lab\_host="SOLR"  
 /clone\_lib="Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Combination of untreated and Phytophthora infestans-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes, tubers, or roots."  
 BASE COUNT 202 a 191 c 144 g 244 t  
 ORIGIN

Query Match 37.9%; Score 655.4; DB 13; Length 781;  
 Best Local Similarity 92.5%; Pred. No. 1.4e-152;  
 Matches 700; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

823 AATGATGCCAATTAATCCGAGCTTATAGTAGTGGAGACCGAGCTGGAAGATATT 882  
 781 AATGATACACATTAATCCGAGCTTATAGTAGTGGAGACCGAGCTGGAAGAT-T 723  
 883 TTCAAGTGGCGGCCATTGATATGCTTATGCTTATCCGATCTGCTCCAGATTT 942  
 722 TTCAACGGCGCCATTGATATGCTTATGCTTATCCGATCTGCTTCCAAATTT 663  
 943 CCTGTGATATTCAGCCATTGAGATATGATGAAGAAAGCGATGAGCTTGGGAAA 1002  
 662 CCTGTGATATTCAGCCATTGAGATATGATGAAGAAAGCGATGAGCTTGGGAAA 603  
 1003 TCCAGATCAAAATCTTGATGAGCTATATCTTATGATGCTGCTGATGTA 1062  
 602 TCCAGATCAAAATCTTGATGAGCTATATCTTATGATGCTGCTGATGTA 543  
 1063 GGATTTGATGATGTTCCAGTTATGGAATTCGACCTGATCAAGCAACAAGAGAGT 1122  
 542 GGATTTGATGATGTTCCAGTTATGGAATTCGACCTGATCAAGCAACAAGAGAGT 483  
 1123 GTATATATGCTGCTTTAGGCTTTAGGCTTGAATCACTAATCAATATCTCAGAT 1182  
 482 GTATATATGCTGCTTTAGGCTTTAGGCTTGAATCACTAATCAATATCTCAGAT 423  
 1183 GTAGAGAAAGATGCGAAGAGAGATATCTGCTCAAGATTTAGCAAGCA 1242  
 422 GTAGAGAAAGATGCGAAGAGAGATATCTGCTCAAGATTTAGCAAGCA 363  
 1243 GGGCTCTCCGAGAGACATATTTGCTGGAAGAGTGAATGAGGAGCACTTTATG 1302  
 362 GGGCTCTCCGAGAGACATATTTGCTGGAAGAGTGAATGAGGAGCACTTTATG 303  
 1303 AAGAAACAAATTCAGAGGCGGAGAAATTTCTTATGAGTCAAGAAAGGTCTCAGAA 1362  
 302 AAGAAACAAATTCAGAGGCGGAGAAATTTCTTATGAGTCAAGAAAGGTCTCAGAA 243  
 1363 CTGAGCTGCTAGAGATGCGCTGTTAATCAGGCTGCTGTTATGCAAGTATTTG 1422  
 242 CTGAGCTGCTAGAGATGCGCTGTTAATCAGGCTGCTGTTATGCAAGTATCTG 183  
 1423 GACGAGATTGAAGCAAGACATCAACAATTCACAAAGAGGCTTATGAGCAAGCA 1482  
 182 GACGAGATTGAAGCAAGACATCAACAATTCACAAAGAGGCTTATGAGCAAGCA 123  
 1483 AAGAACTTCTACCTTGCCATTGCTTATGCAAAATCTTGTGCCCCCTTATAGAACT 1542  
 122 AAGAACTTCTACCTTGCCATTGCTTATGCAAAATCTTGTGCCCCCTTATAGCAACT 63

QY 1543 TCCTCTCCACTAGCAAAAGACATGATGATGTA 1579  
 Db 62 TCCTCTCCACTAGCAAAAGACATGATGATGTA 26

RESULT 2  
 B0046203  
 LOCUS  
 DEFINITION  
 B0046203 751 bp mRNA linear EST 10-MAR-2003  
 EST595321 P. infestans-challenged potato leaf, incompatible  
 reaction Solanum tuberosum cDNA clone BPL14E21 5' end, mRNA  
 sequence.

ACCESSION  
 B0046203  
 VERSION  
 B0046203.1 GI:19820189  
 KEYWORDS  
 SOURCE  
 ORGANISM

Solanum tuberosum (potato)  
 Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanales; Solanales; Solanaceae; Solanum.  
 1 (bases 1 to 751)  
 Zhang, P., Hernandez, M., Tornqvist, C.-E., Wirtz, U., Loukianov, A.,  
 Rangel, P., Haberlach, G.T., Karayancheva, S.A., Tsai, J., Chieningo, A.,  
 Bougri, O., Buell, C.R., Romming, C.M., Helgeson, J. and Baker, B.  
 Generation of ESTs from Potato Leaves Challenged with Phytophthora  
 infestans, incompatible interaction (2002)  
 Unpublished

# JOURNAL

Contact: Robin Buell

The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

This clone can be obtained from the University of Arizona Genomics  
 Institute. Orders can be made through URL:  
<http://genome.arizona.edu/orders/>

Seq primer: T3.

# FEATURES

Location/Qualifiers

1..751  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="BPL14E21"  
 /tissue\_type="leaf"  
 /dev\_stage="6 week old"  
 /lab\_host="SOLR"  
 /clone\_lib="P. infestans-challenged potato leaf,  
 incompatible reaction"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; supplier: UC Berkeley, pGEC; sequencing: The  
 Institute for Genomic Research. Whole plants were  
 challenged with 450,000 sporangia/ml P. infestans isolate  
 US-1 (US940501) in Bioron (Medison, Wisconsin). Leaf  
 tissue was collected at 1, 2, 5, 12, and 24 hours  
 post-challenge and frozen in liquid nitrogen immediately  
 upon removal. Kennebec plants showed no signs of HR.  
 Katchadin plants (susceptible to P. infestans US-1) were  
 used as controls and showed infection. NOTE: We cannot  
 exclude the possibility that this sequence is actually  
 derived from Phytophthora rather than potato."

# BASE COUNT

219 a 119 c 199 g 214 t

Query Match 36.0%; Score 621.6; DB 12; Length 751;  
 Best Local Similarity 91.8%; Pred. No. 3.9e-144;  
 Matches 690; Conservative 0; Mismatches 59; Indels 3; Gaps 3;

QY 438 AATGCTGATCAAGATATTCATGCTTGGTGATCAAGACTGAAGAGGAGCACTTT 497  
 Db 1 AATGCTGATCAAGATATTCATGCTTGGTGATCAAGACTGAAGAGGAGCACTTT 60  
 QY 498 CTCGTACAGTCCAGTTTGGTGGTACCCAGCTGGAAGAAATGCTGTGTATATAGAA 557  
 Db 61 TTCGTACAGTCCAGTTTGGTGGTACCCAGCTGGAAGAAATGCTGTGTATATAGAA 120



QY	558	AAAGTGATGATGTCGATATTAACAGACGACCTTTAGTGAAGAGACAGCTGAGATCTAC	617
Db	121	AAAGGTGATGAGGTGCTATTTAAGACAGGACCTTTAGTGAAGAGACATCTATATCTAC	180
QY	618	CGATGATTTAGAAAGTGAAGCCGGATATTTGTTTCAGAGGAATTTGGCTTTGTAAGTGA	677
Db	181	TGAGAGCATAGAAAGTGAAGCCGGATATTTGTTTCAGAGGAATTTGGCTTTGTAAGTGA	240
QY	678	AGCATATGATTCGTTGTGGCGAAAGTATGTGCAGAGTATGTCAAAAGACATTTTACTTTGGAAC	737
Db	241	AGCATATGATTCGTTGTGGCGAAAGTATGTGCAGAGTATGTCAAAAGACATTTTACTTTGGAAC	300
QY	728	CAAGCTAATGACCCAGAGAGAAAGAGCTATCTGGGCAATATATGTGTGTGACAGAG	797
Db	301	CATGCTAATGATCTCCAGACAGAAAGAGCTATCTGGGCAATATATGTGTGTGACAGAG	360
QY	798	AACGATGAGCTTGTGTGATGAGCCCTTAATGATCCACATTAATCCGCAAGCTTTAGATAG	857
Db	361	AACGATGAGCTTGTGTGATGAGCCCTTAATGATCCACATTAATCCGCAAGCTTTAGATAG	420
QY	858	GTGGAGAGACAGGCTTGGAAGATATTTTCAGTGGCGGCGCATTTGATATGCTTGATGCTGC	917
Db	421	GTGGAGAGGCGAGGCTTGGAAGATATTTTCACCGGGCGCATTTGATATGCTTGATGCGAGC	480
QY	918	TTTATCCGATCTGCTCTCCAGATTTCCCTGTTATATTCAGCATTTAGAGATATGATTTGA	977
Db	481	TTTATCCGATCTGCTCTCCAGATTTCCCTGTTATATTCAGCATTTAGAGATATGATTTGA	540
QY	978	AGGAATGCGTATGACCTTGTGGAATCCAGATACAAAATTTTCGATGAGCTATATCTCTA	1037
Db	541	AGGAATGCGTATGACCTTGTGGAATCCAGATACAAAATTTTCGATGAGCTATATCTCTA	600
QY	1038	TTGTTACTATGTTGCTGTGATCTGTAGG-ATTGATGAGTGTTCAGTATATGGTATTTGCAC	1098
Db	601	TTGTTACTATGTTGCTGTGATCTGTAGG-ATTGATGAGTGTTCAGTATATGGTATTTGCAC	660
QY	1097	CTGATCAAAAGGCAACAAGAGAGTGATATATGATGCTGC-TTTGGCTTTAGGGCTTGCA	1155
Db	661	CTGATCAAAAGGCAACAAGAGAGTGATATATGATGCTGC-TTTGGCTTTAGGGCTTGCA	720
QY	1156	AATCAACTPACCAATATATCTCAGAGATGTAGG	1187
Db	721	AATCAACTPACCA-ATATACTCAAAAGATGTAGG	751
RESULT 3			
LOCUS	BM410846	787 bp	mRNA linear EST 22-JAN-2002
DEFINITION	EST585173 tomato breaker fruit Lycopersicon esculentum cDNA clone		
ACCESSION	BM410846		
VERSION	BM410846.1	GI:18262476	
KEYWORDS	EST.		
SOURCE	Lycopersicon esculentum (tomato)		
ORGANISM	Lycopersicon esculentum		
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (baes 1 to 787)		
AUTHORS	Alcala,J., Vebelav,J., White,R., Vision,T., Karanmycheva,S.A., Tsai,J., Bougari,O., Kirkness,E., Uteback,T., Van Aken,S., Roming,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.		
TITLE	Generation of ESTs from tomato fruit tissue, breaker stage (2002)		
JOURNAL	Unpublished		
COMMENT	Contact: CUGI Clemson University Genomics Institute 100 Jordan Hall, Clemson, SC 29634, USA Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a> This clone is available through the Clemson University Genomics Institute Seq primer: T3.		

FEATURES	SOURCE
location/Qualifiers	1. .787
/organism="Lycopersicon esculentum"	
/mol_type="mRNA"	
/culivar="TA496"	
/db_xref="taxon:4081"	
/clone="cLEG54F18"	
/tissue_type="pericarp"	
/dev_stage="breaker"	
/lab_host="SOLR"	
/clone_id="tomato breaker fruit"	
/note="Vector: pBluescriptSKmcutap; Site_1: EcoRI; Site_2: XhoI; supplier: Boyce Thompson Institute; Fruit were harvested at the breaker stage (first sign of lycopen accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."	230 a 132 c 211 g 214 t

Query Match	35.0%;	Score 605.2;	DB 12;	Length 787;
Best Local Similarity	85.6%;	Pred. No. 5e-140;		
Matches 673;	Conservative 0;	Mismatches 113;	Indels 0;	Gaps 0;

Qy	491	GCACCTTTCCTGTAACAGTCCAGTTGGGTGATGCCAGCTGGAGAAATGACTGTGCAT	550
Db	2	GGAATTTTCTGTACGGTCTGCTATTTTGGTACTCCATCTGGAGAAOGGACGATGACAT	61
Qy	551	CAGAGAAAAAGGTATGATGTGTATTTAAACGAGCGACTTTAGTGAAGGCACTGA	610
Db	62	CGGAACAGATGGTCTTATGATGTGTATTTGAGGCGAGGCGCTTGATAGAGGCAACTGA	121
Qy	611	GATCTACCGATGATTTTGAAGTGAAGCCGGAATATGTGTCCAGGGAAATTGGGCTGT	670
Db	122	GATTTACCAATGAGTTTGAAGTGAAGCCGGAATATCTATTTCCGGGAAATTTGGGCTTGT	181
Qy	671	TGAGTGAAGCATATGATCGTTGTGGCCGAGATGTGCAAGTATGCAAGACATTTTACT	730
Db	182	TGAGTGAAGCATATGATGAGTGTGTGAGTATGTGCAAGTATGCAAGACGTTTAACT	241
Qy	731	TAGGAACCAAGCTTAATGACCCGAGAGAGAAAGAAAGATCTGGGCAATTAATGTGGT	790
Db	242	TAGGAACCTATGCTATATATACCTCCGAGAGAAAGAAAGGCTATCTGGGCAATTAATGTAGT	301
Qy	791	GCAAGAGACGSAATGAGCTGTGTGATGAGCCCTTAATGCATCCACATTAATCCGCAAGCTT	850
Db	302	GCAGAAAGACAGATGAAGCTTGTGATGGCCAAACGATCATATATTACCOCGCAAGCTT	361
Qy	851	TAGATAGGTGGAGACCAAGCTGGAAGATATTTCACTGGGCGGCCATTGTATATGCTTG	910
Db	362	TAGATAGGTGGAGAAATAGGCTGAGAAATGTTTTCAATATGGCGCGCAATTTGACATGCTCG	421
Qy	911	ATGCTGCTTATATCCGATACTGTCTCCAGATTTTCCGTTGATATTCAGCAATTCAGAGATA	970
Db	422	ATGCTGCTTGTTCGGAATACAGCTTTCTAACTTTCCAGTTGATATTACGCAATTCAGAGATA	481
Qy	971	TGATTTGAAGAAATCGTATGAGCTTTGTGAAATCCAGATACAAACCTTTCGATGAGCTAT	1030
Db	482	TGATTTGAAGAAATCGTATGAGCTTTGAGAAATCGATATCAAAACCTTTCGACAACTAT	541
Qy	1031	ATCTCTATGTTACTATGTTGTGCTGTGCTGTAGAGATTGATGAGAGTTCGAGTTATGGGTA	1090
Db	542	ACCTTTATGTTATTAATGTTGTGCTGTGCTGTGAGAGTTCGAGTTATGGGTA	601
Qy	1091	TTGCACTGTGATCAAGAGCAACAACAGAGAGTGTATATAATGCTCTTTGGCTTTAGGCG	1150
Db	602	TCGCGCCTGTGATCAAGAGCAACAACAGAGAGCGTATTAATGCTCTTTGGCTTCGGGGA	661
Qy	1151	TTTGAATTCAACTTACCAATATATCTCAGATATGAGAGAAAGTCCAGAAAGAGAAAG	1210
Db	662	TCGGAATTCAACTTACCTTACCAATATCTCAGATATGTTGAGAGAAAGTCCAGAAAGAGAAAG	721

QY	1211	TATATCTGCGCCCAAGATGAATTTAGACAGGAGGCGGCTCTCCGAGCAAGACATATTGGCTG	1270
DB	722	TTCTATCTTCCCTCCAGATGAATTTAGACAGGAGGCGGCTCTATCCGATGAAGATATTATTGGCTG	781
QY	1271	GAAGAG	1276
DB	782	GAGGG	787
RESULT 4	BG351357	677 bp	mRNA linear EST 11-SEP-2002
LOCUS	104A12	Mature tuber lambda ZAP	Solanum tuberosum cDNA, mRNA
DEFINITION	sequence.		
ACCESSION	BG351357		
VERSION	BG351357.1	GI:13180099	
KEYWORDS	EST.		
SOURCE	Solanum tuberosum (potato)		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.		
REFERENCE	1 (bases 1 to 677)		
AUTHORS	Crookshanks,M., Emmeren,J., Wellinder,K.G. and Nielsen,K.L.		
TITLE	The potato tuber transcriptome: analysis of 6077 expressed sequences		
JOURNAL	FEBS Lett. 506 (2), 123-126 (2001)		
MEDLINE	21475600		
PUBMED	11591384		
COMMENT	Contact: Karen G. Wellinder Institut for bioteknologi Aalborg Universitet Sohnsgaardsholmsvej 49, 9000 Aalborg, Denmark Tel: +45 96358467 Fax: +45 98141808 Email: kgw@bio.auc.dk Sequenced from the 5' end. High quality sequence scop: 677 POLYA=No.		
FEATURES	Location/Qualifiers		
source	1..677		
	/organism="Solanum tuberosum"		
	/mol_type="mRNA"		
	/cultivar="Field grown Kuras"		
	/db_xref="taxon:4113"		
	/tissue_type="tuber"		
	/clone_id="Mature tuber lambda ZAP"		
	/note="Vector: Lambda ZAP"		
BASE COUNT	192 a 102 c 198 g 185 t		
ORIGIN			
Query Match	32.9%; Score 568.4; DB 10; Length 677;		
Best Local Similarity	90.9%; Pred. No. 7.4e-131;		
Matches	616; Conservative 0; Mismatches 61; Indels 1; Gaps 1.		
QY	316	GATTGAGTCCGAGAAACCGCGCTCTTGATATCATCCAGGTTCTTACTGCTGAGATAGG	375
DB	1	GATTGAGTCCGAGAAACCGCGCTCTTGATATCATCCAGGTTCTTACTGCTGAGATAGG	59
QY	376	AATTGAGTGGAAAGGGAGATCAAGAAAGGTGGAGACAAGGTGAATTTGGGCTCT	435
DB	60	AATTGAGTGGAAAGGGAGATCAAGAAAGGTGGAGACAAGGTGAATTTGGGCTCT	119
QY	436	TTAATTTGTCGATCCAAAGATATTCATGCTTGGGTGATCAAGAACTGAAAGGAGCACT	495
DB	120	TTAATTTGTCGATCCAAAGATATTCATGCTTGGGTGATCAAGAACTGAAAGGAGCACT	179
QY	496	TTCTCTGTACAGTCCAGTTGGTGGCTAGCCAGCTGAGAAATGACTGTGTATCATCAGG	555
DB	180	TTCTCTGTACAGTCCAGTTGGTGGCTAGCTGAGAAATGACTGTGTATCATCAGG	239
QY	556	AAAAAGGTGTATGATGTGTATTTAAAGCAGGAGCTTTTAGTAAAGGACGTGAGATCT	615

FEATURES	source
Db	AAAAAGGTGATGAGGTGATTGTGAACGAGCACCTTTAGTGAAGAGGCATCTGATATCT 229
Qy	240
Db	616 ACCGATGATTTAGAAAGTGAAGCCGCGATATTGTTGTTCCAGGAAATTTGGCTTGTAGT 675
Qy	300 ACTGAGGACATAGAAAGTGAAGCCCGAATATTGTTGTTCCGGTAAATTTGGCTTGTAGT 359
Qy	676 GAAGCATATGATCGTTGTGGCGAAGTATGTCAGAGTATGCAAGACATTTTACTTAGA 735
Db	360 GAAGCATATGATCGTTGTGGCGAAGTATGTCAGAGTATGCTATAGACATTTTACTTAGA 419
Qy	736 ACCAAGCTAATGACCCCGAAGAGAAGAAAGACTATCTGGCGAATATATGTGTGTCAGG 795
Db	420 ACCATGCTAATGATCTCCGACAGAAAGAGACTATCTGGCGAATATATGTGTGTCAGG 479
Qy	796 AGAACGATGATGCTTTGTGATGTCGCCCTTATGTCATCCCATATCTCCGCAAGCTTTAGT 855
Db	480 AGAATGATGATGCTTTGTGATGTCGCCCTTATGTCATCCCATATCTCCCAAGCTTTAGT 539
Qy	856 AGGAGGAGACACGAGCTGGAAGATATTTTCAGTGGGGGGCCCATTTGATGATGCTGATGCT 915
Db	540 AGGAGGAGGAGCAGAGCTGGAAGATATTTTCACGGGGGGCCCATATGATGCTGATGCA 599
Qy	916 GCTTATTCGATGATGCTGCTCCAGATTTTCCTGTGATATTCAGCCATTCAGAGATATGAT 975
Db	600 GCTTATTCGATGATGCTTTCGAATTTTCCTGTGATATTCAGCCATTCAGAGATATGAT 659
Qy	976 GAAGCAATGCTGATGAC 993
Db	660 GAAGCAATGCTGATGAC 677
RESULT 5	
LOCUS	AM442101 686 bp mRNA linear EST 18-MAY-2001
DEFINITION	ESTJ311497 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA clone cLEN21G20 5', mRNA sequence.
ACCESSION	AM442101
VERSION	AM442101.1 GI:6977352
KEYWORDS	EST.
SOURCE	Lycopersicon esculentum (tomato)
ORGANISM	Lycopersicon esculentum
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 686)
AUTHORS	Alcala,J., Yrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Romling,C.M., Fraser,C.M., Martin,G.B., Tankley,S.D. and Giovannoni,J.
TITLE	Generation of ESTs from tomato fruit tissue
JOURNAL	Unpublished
COMMENT	Contact: CUGI Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a> 5 prime sequence. Location/Qualifiers 1..686 /organism="Lycopersicon esculentum" /mol_type="mRNA" /cultivar="T496" /db_xref="taxon:4081" /clone="cLEN21G20" /tissue_type="pericarp" /dev_stage="red ripe (7-20 days post-breaker)" /clone_lib="tomato fruit red ripe, TAMU" /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Giovannoni; Fruit were tagged at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit) and harvested 7 days post-breaker (fully red-ripe), 10 days post breaker, and 20 days post-breaker (over-ripe). 20 day fruit which showed external or internal signs of pathogenesis were

discarded. Fruit were cut in half and the seeds and  
locules were discarded prior to freezing the pericarp."

BASE COUNT 204 a 114 c 182 g 186 t

ORIGIN

Query Match 32.0%; Score 552.2; DB 9; Length 686;  
Best Local Similarity 87.9%; Pred. No. 8.3e-127;  
Matches 602; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 565 TATGATGATGATTAAGACAGCGCTTTAGTGAAGAGGAGCGTGAATCTACCGATGAT 624  
DB 2 TATGATGATGATTTAGGAGGAGCGCTTTAGTGAAGAGGAGCGTGAATCTACCGATGAT 61  
QY 625 TTAGAAGTGAAGCGCGATATTTGTTTCAAGGAAATTTGGCTTTGAGTGAAGATAT 684  
DB 62 TTAGAAGTGAAGCGCGATATTTGTTTCAAGGAAATTTGGCTTTGAGTGAAGATAT 121  
QY 685 GATCGTTGTTGGCGAAGATGTCAGAGTATGCAAGACATTTTACCTTAGAACCAAGCTA 744  
DB 122 GATAGGTGTGTGAAGTATGTGAGAGTATGCAAGACATTTTACCTTAGAACCAAGCTA 181  
QY 745 ATAGCCCGAGAGAGAGAGAGAGTATGCGGCAATATATGTTGTCAGAGAGAGAGAGAT 804  
DB 182 ATAGCTCCGAGAGAGAGAGAGAGTATGCGGCAATATATGTTGTCAGAGAGAGAGAT 241  
QY 805 GAGCTTTGTTGATGCGCTTAATGATCCGACATTAATCTCCGACGTTAGATAGTGGAG 864  
DB 242 GAACTTTGTTGATGCGGCAAGAGATATATTAATCCCGGACGCTTAGATAGTGGAG 301  
QY 865 ACCAGGCTGAAGATATTTTCAAGTGGCGGCGCAATTTGATATGCTGATGCTTTATCC 924  
DB 302 AATAGGCTGAAGATATTTTCAAGTGGCGGCGCAATTTGATATGCTGATGCTTTATCC 361  
QY 925 GATACGTCTCCGATTTCTGTTGATATCAGCAATTCAGATATGATATGATGAAGAGAT 984  
DB 362 GATACGTCTCTAATCTTTCAATTTGATATCAGCAATTCAGATATGATGAAGAGAT 421  
QY 985 CGTATGAGCTTTGGAATCCAGATCAAAATCTTGCATGAGCTATATCTATTTGATC 1044  
DB 422 CGTATGAGCTTTGGAATCCAGATCAAAATCTTGCATGAGCTATATCTATTTGATC 481  
QY 1045 TATGTTGCTGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1104  
DB 482 TATGTTGCTGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541  
QY 1105 AAGGCAACACAGAGATGATATATATGCTTTGCTTTAGGCTTGAATCAACTA 1164  
DB 542 AAGGCAACACAGAGATGATATATATGCTTTGCTTTAGGCTTGAATCAACTA 601  
QY 1165 ACCAATATATCTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1224  
DB 602 ACTACACTACTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661  
QY 1225 GATGAATTAGCAGAGCGGCTCT 1249  
DB 662 GATGAATTAGCAGAGCGGCTCTAT 686

RESULT 6  
BM408984 765 bp mRNA linear EST 22-JAN-2002  
LOCUS EST563311 tomato breaker fruit Lycopersicon esculentum cDNA clone  
DEFINITION CLE646P23 5' end, mRNA sequence.  
ACCESSION BM408984  
VERSION BM408984.1 GI:18260614  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Bukarjora; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asterales; Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
REFERENCE 1 (bases 1 to 765)  
Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Tsai

J., Bougri,O., Kirkness,E., Uterback,T., Van Aken,S., Roming  
,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.  
Generation of ESTs from tomato fruit tissue, breaker stage (2002)  
Unpublished

COMMENT

CONTACT: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

This clone is available through the Clemson University Genomics

Institute

Seq primer: T3

#### FEATURES

source

1..765

/organism="Lycopersicon esculentum"

/mol\_type="mRNA"

/culivar="TA496"

/db\_xref="taxon:4081"

/clone="CLE646P23"

/tissue\_type="Pericarp"

/dev\_stage="breaker"

/lab\_host="SOLR"

/clone\_lib="tomato breaker fruit"

/note="Vector: pBluescriptSMK2adapt; Site 1: EcoRI;  
Site 2: XhoI; supplier: Boyce Thompson Institute;  
sequencing: The Institute for Genomic Research. Fruit  
were harvested at the breaker stage (first sign of  
lycopene accumulation on the blossom end of fruit). Fruit  
were cut in half and the seeds and locules were discarded  
prior to freezing the pericarp."

BASE COUNT 237 a 122 c 202 g 204 t

ORIGIN

Query Match 31.9%; Score 551.2; DB 12; Length 765;  
Best Local Similarity 86.1%; Pred. No. 1.5e-126;  
Matches 634; Conservative 0; Mismatches 98; Indels 4; Gaps 2;

QY 625 TTAGAAGTGAAGCGGATATTTGTTTCAAGGAAATTTGGCTTTGAGTGAAGATAT 684  
DB 1 TTAGAAGTGAAGCGGATATTTGTTTCAAGGAAATTTGGCTTTGAGTGAAGATAT 60  
QY 685 GATCGTTGTTGGCGAAGATGTCAGAGTATGCAAGACATTTTACTTAGAACCAAGCTA 744  
DB 61 GATAGGTGTGTGAAGTATGTGAGAGTATGCAAGACATTTTACTTAGAACCAAGCTA 120  
QY 745 ATGACCCCGAGAGAGAGAGAGCTATCTGGCAATATATGTTGTTGTCAGAGAGAGAT 804  
DB 121 ATGACCTCCGAGAGAGAGAGAGCTATCTGGCAATATATGTTGTTGTCAGAGAGAGAT 180  
QY 805 GAGCTTTGTTGATGCGCTTAATGATCCGACATTAATCTCCGACGTTAGATAGTGGAG 864  
DB 181 GAACTTTGTTGATGCGGCAAGAGATATATTAATCCCGGACGCTTAGATAGTGGAG 240  
QY 865 ACCAGGCTGAAGATATTTTCAAGTGGCGGCGCAATTTGATATGCTGATGCTTTATCC 924  
DB 241 AATAGGCTGAAGAGATGTTTCAATGGGCGCATTTGAGATGCTGATGCTTTGTC 300  
QY 925 GATCTGTCTCAAGATTTCTGTTGATATTCAGCAATTCAGAGATGATGATGATGATGAT 984  
DB 301 GATACAGTTTCTAATCTTTCAATTTGATATTCAGCAATTCAGAGATGATGATGATGAT 360  
QY 985 CGTATGAGCTTTGGAATCCAGATCAAAATCTTGCATGAGCTATATCTATTTGATC 1044  
DB 361 CGTATGAGCTTTGGAATCCAGATCAAAATCTTGCATGAGCTATATCTATTTGATC 420  
QY 1045 TATGTTGCTGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1104  
DB 421 TATGTTGCTGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
QY 1105 AAGGCAACACAGAGATGATATATATGCTTTGCTTTAGGCTTGAATCAACTA 1164  
DB 481 AAGGCAACACAGAGATGATATATATGCTTTGCTTTAGGCTTGAATCAACTA 540

Oy		1165	ACCAATATATCTCAGGATGTGGAGAATGCGACAAGAGGAAGTGTACTTGCTCAT	1224
Dd		541	ACTACATATCTCAGATGATTGGAAAGATGCCAAGAAGAGTCTACTTGCTTCAA	600
Oy		1225	GATGATTTAGGACACAGCAGGGCTCTCCGCAGAGACATA-TTTCGTGAGAGTAGCTGA	12833
Dd		601	GATGATTTAGGACACAGCAGGCTCTATCCGATGAGATATATTTTGGTGGAAAGGTGACC	660
Oy		1284	TAACTGGAGGAACCTTTATGAGAAAACAATTGAGAGGGCGAGGAAATCTTTGATGATC	13433
Dd		661	TAAAGGAGGAATCTTTATGAGAAAACA---TACTAGGGCAAGAAAGTCTTGATGAGACA	717
Oy		1344	AGAGAAAGGTGCACA	1359
Dd		718	GAGAAAGGCTGACAGA	733
RESULT 7				
LOCUS		BM409200	682 bp	mRNA linear EST 22-JAN-2002
DEFINITION		EST5835327 tomato breaker fruit Lycopersicon esculentum cDNA clone		
ACCESSION		BM409200		
VERSION		BM409200.1	GI:18260830	
KEYWORDS		EST.		
SOURCE		Lycopersicon esculentum (tomato)		
ORGANISM		Lycopersicon esculentum		
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 682)		
AUTHORS		Alcala,J., Vrebalov,J., White,R., Vision,T., Karamychava,S.A., Tsai,J., Bougri,O., Kirtnes,E., Utterback,T., Van Aken,S., Roming,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.		
JOURNAL COMMENT		Generation of ESTs from tomato fruit tissue, Breaker stage (2002) Unpublished Contact: CUGI Clemson University Genomics Institute 100 Jordan Hall, Clemson, SC 29634, USA Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a> This clone is available through the Clemson University Genomics Institute Seq primer: T3. Location/Qualifiers 1..682 /organism="Lycopersicon esculentum" /mol_type="mRNA" /cultivar="TA496" /db_xref="taxon:4081" /clone="CLBG47J1" /tissue_type="pericarp" /dev_stage="breaker" /lab_host="SOLR" /note="Vector: pBluescriptSmCladap; Site 1: EcoRI; Site 2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopen accumulation on the blossom end of fruit). Fruits were cut in half and the seeds and locules were discarded prior to freezing the pericarp."		
BASE COUNT		217 a	114 c	173 g
ORIGIN				178 t
Query Match		31.3%	Score 541.2,	DB 12, Length 682;
Best Local Similarity		87.1%;	Pred. No. 4,6e-124;	
Matches 594;		Conservative 0;	Mismatches 88;	Indels 0; Gaps 0;
Oy		706	GCAGAGTGTGCAAAACATTTTACTTAGAACCAACGTAATGACCCCGAGAGAAAGA	765
Dd		1	GAGAGTATGCAAAACGTTTAACTTAGAACCTAATGATCTCCGAGAGAAAGG	60

FEATURES	LOCATION/Qualifiers	1..672
Db	61	GCATCTGCGGCAATATATATATGTGTGGTGCAGAGAAACGATGAGCTTGTGATGAGCCCTAT
Qy	826	GCATCCCAACATAACTCCGCAAGCTTTAGATAGTGGAGACACAGCTGGAAAGATATTTTC
Db	121	GCATCATATATATTAACCCGGGAGCTTGAATAGTGGGAAAATATGAGCTGAAGATGTTTTTC
Qy	886	AGTGGGCGGCGCATTTGATATGCTTGATGCTGCTTTATCCGATACTGTCTCCAGATTTCCCT
Db	181	AATGGGGCGGCATTATGACATGCTCGATGAGTGCTTTGTCGATVACAGTTCTTAACCTTTC
Qy	946	GTTATATATTCAGCCATTTCAGAGATATGATTTAGAGAAATGCGTATGACCTTGTGAAATCC
Db	241	GTTGATATTCAGCCATTTCAGAGATATGATTTAGAGAAATGCGTATGACCTTGTGAAAAATCG
Qy	1006	AGATPACAAAACCTTTGATGACCTATATATCTATTTGTTATCTATGTTGCTGTACTGTAGCA
Db	301	AGATPACAAAACCTTCGACGAACCTATACCTTTATTTGTTATTTATGTTGCTGTATCGGATGCG
Qy	1066	TTGATATGATGTTCCAGTTATGCGATATGACCTTAATCAAGGCAACAGAGAGTGTA
Db	361	TTGATATGATGTTCCAAATTAATGAGATGCGCCCTTAATCAAGGCAACAGAGAGTGTA
Qy	1126	TATATATGCTGTTTGGCTTTAGGCTTTCGCAATCAATCAACCAATATATCTCAGAGATGTA
Db	421	TATATATGCTGTTTGGCTTGGCTTGGGATGCGAAATCAATTAATCAATATCTCAGAGATGTT
Qy	1186	GGAGAGAGATGCCAGAGAGAGAGAGATATCTTGCTCTCAGATGATTAATGCAACAGGACGG
Db	481	GGAGAGAGATGCCAGAGAGAGAGAGATATCTTGCTCTCAGATGATTAATGCAACAGGACG
Qy	1246	CTCTCCGACGAGACATATTTGCTGGAAGATGACTGATATGATGAGAGAACTTTATGTAAG
Db	541	CTATTCGATGAGATATATTTTCTGGAAGGATGACCGATTAATGAGAAATCTTTATGAG
Qy	1306	AAACAAATTCAGAGGGCGAGAGAAATCTTTGATGATGATCAGAGAAAGGTGTACAGAACTG
Db	601	AAACAAATATACATAGGGCAAGAAAGCTTTGATGATGAGCAGAGAAAGGGGTGACAGAAATG
Qy	1366	GACTCTGCTAGTATGATGAGCTG 1387
Db	661	AGCTCAGCTAGTATGATTCCTG 682
RESULT 8		
AM442407		672 bp mRNA linear EST 18-MAY-2001
LOCUS	ESTJ311803 tomato fruit red ripe, TMAU Lycopersicon esculentum cDNA	
DEFINITION	clone GLEN22P14 5', mRNA sequence.	
ACCESSION	AM442407	
VERSION	AM442407.1 GI:6977658	
KEYWORDS	EST.	
SOURCE	Lycopersicon esculentum (tomato)	
ORGANISM	Lycopersicon esculentum	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.	
AUTHORS	1 (bases 1 to 672)	
TITLE	Alcala,J., Vredalov,J., white,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Rönning,C.M., Fraeier,C.M., Martin,G.B., Tankeley,S.D. and Giovannoni,J.	
JOURNAL	Generation of ESTs from tomato fruit tissue	
COMMENT	Unpublished	
CONTACT	Contact: CUGI	
CLUSTER	Clemson University Genomics Institute	
ADDRESS	Clemson University	
LOCATION	100 Jordan Hall, Clemson, SC 29634, USA	
EMAIL	Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a>	
SEQUENCE	5 prime sequence.	

```

/mol_type="mRNA"
/cultivar="7496"
/db_xref="taxon:4081"
/clone="cLEB2L14"
/tissue_type="pericarp"
/dev_stage="red ripe (7-20 days post-breaker)"
/clone_lib="tomato fruit red ripe, TANO"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Giovannoni; Fruit were tagged at the
breaker stage (first sign of lycopen accumulation on the
blossom end of the fruit) and harvested 7 days
post-breaker (fully red-ripe), 10 days post breaker, and
20 days post-breaker (over-ripe). 20 day fruit which
showed external or internal signs of pathogenesis were
discarded. Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."

```

BASE COUNT 212 a 112 c 170 g 178 t

ORIGIN

Query Match 30.8%; Score 532.8; DB 9; Length 672;  
 Best Local Similarity 87.1%; Pred. No. 5.8e-122;  
 Matches 585; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

```

719 AGACATTTTACTTGAAGACCAAGCTAATGACCCGAGAGAAAGAGCTATCTGGGCAA 778
1 AGACCTTTAACTAGGAACCTATGATGACTCCGAGAGAAAGGCTATCTGGGCAA 60
779 TATATGTGTGTGCGAGAGAGAGAGCTGTGTGATGCGCTTAATGATCCCACTA 838
61 TATATGTATGTGTGCGAGAGAGAGAGCTGTGTGATGCGCTTAATGATCCCACTA 120
839 CTCCGCAAGCTTTAGATAGTGGGAGACGAGCTGGAGATATTTTCACTGGGCGCCAT 898
121 CCCCGGAGCCTTATAGTGTGGGAAATAGGCTGAAATATTTTCACTGGGCGCCAT 180
899 TTGATATGCTTATGCTGCTTATTCGATATCTGCTCCAGATTTTCTGTGATATTCAC 958
181 TTGATATGCTTATGCTGCTTATTCGATATCTGCTCCAGATTTTCTGTGATATTCAC 240
959 CATTGAGAGATGATGATGAGAGAGAGAGCTGTGTGATGCGCTTAATGATCCCACTA 1018
241 CATTGAGAGATGATGAGAGAGAGAGCTGTGTGATGCGCTTAATGATCCCACTA 300
1019 TCGATGAGCTATATCTTATGCTTATGCTGCTTATGCTGCTTATGCTGCTTATGCTGCTT 1078
301 TCGATGAGCTATATCTTATGCTTATGCTGCTTATGCTGCTTATGCTGCTTATGCTGCTT 360
1079 CAGTTATGGGATGTCACCTGATCAAGAGCAACAGAGAGCTATATGCTGCTT 1138
361 CAGTTATGGGATGTCACCTGATCAAGAGCAACAGAGAGCTATATGCTGCTT 420
1139 TGGCTTTAGGGCTTGAATCAATCAATATATCTGAGATGTAAGAGAGAGAGAGAGAGAG 1198
421 TGGCTTTAGGGCTTGAATCAATCAATATATCTGAGATGTAAGAGAGAGAGAGAGAGAG 480
1199 GAAGAGAGAGATATCTGCTCAAGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1258
481 GAAGAGAGAGATATCTGCTCAAGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
1259 ACATATTTGCTGGAAGAGTACTGATTAAGTGAAGAACTTTTATGAAGAACTTATGAGA 1318
541 ACATATTTGCTGGAAGAGTACTGATTAAGTGAAGAACTTTTATGAAGAACTTATGAGA 600
1319 GGGGAGAGAAATTTCTTATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1378
601 GGGGAGAGAAATTTCTTATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
1379 GATGGCTGTGT 1390
661 GATTCCTGTAT 672

```

RESULT 9

```

BE433198 642 bp mRNA linear EST 18-MAY-2001
LOCUS BE433198
DEFINITION EST399727 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
ACCESSION BE433198
VERSION BE433198.1
KEYWORDS GI:9431041
SOURCE EST
ORGANISM Lycopersicon esculentum (tomato)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 642)
Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Niernman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
,S.D.
Generation of ESTs from tomato fruit tissue, breaker stage
JOURNAL Unpublished
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
source
1..642
location/Qualifiers
1..642
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="7496"
/db_xref="taxon:4081"
/clone="cLEB12N5"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOL"
/clone_lib="tomato breaker fruit, TIGR"
/notes="Vector: pBluescriptSMCtdapt; Site 1: EcoRI;
Site 2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopen accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."
BASE COUNT 194 a 103 c 167 g 178 t
ORIGIN
Query Match 30.2%; Score 522; DB 10; Length 642;
Best Local Similarity 88.3%; Pred. No. 2.8e-119;
Matches 567; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
660 TTTGGGCTTTGAGTGAAGATATGATGCTTGTGCGAAGATATGCGAGATATGCAAA 719
1 TTTGGGCTTTGAGTGAAGATATGATGCTTGTGCGAAGATATGCGAGATATGCAAA 60
720 GACATTTTACTTGAAGACCAAGCTATATGACCCGAGAGAGAGAGAGAGAGATCTGGGCAT 779
61 GACATTTTACTTGAAGACCAAGCTATATGATGCTTGTGCGAAGATATGCGAGATATGCAAA 120
780 ATATGCTGTGTGCGAGAGAGAGAGATGATGCTTGTGCGAAGATATGCGAGATATGCAAA 839
121 ATATGCTGTGTGCGAGAGAGAGAGATGATGCTTGTGCGAAGATATGCGAGATATGCAAA 180
840 TCCGCAAGCTTTAGATAGTGGAGAGACGAGCTGGAAGATATTTTCACTGGGCGCCAT 899
181 CCCGCAAGCTTTAGATAGTGGAGAGACGAGCTGGAAGATATTTTCACTGGGCGCCAT 240
900 TGAATGCTTATGCTGCTTATTCGATATGCTGCTTATTCGATATGCTGCTTATTCGATAT 959
241 TGAATGCTTATGCTGCTTATTCGATATGCTGCTTATTCGATATGCTGCTTATTCGATAT 300
960 ATTGAGATATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1019
301 ATTGAGATATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
1020 CGATGAGCTATATCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 1079

```

Db 361 CGAGCAATATACCTTTATTTGTTATTTGTTGCTGTGATCGTTGGCTGATGAGTGTCC 420  
 Qy 1080 AGTATGGGTATGTGACCTGATCAAGGCAACACAGAGTGTATTTATGCTGCTT 1139  
 Db 421 AATTATGGGTATGCGCCCTGATCAAGGCAACACAGAGTGTATTTATGCTGCTT 480  
 Qy 1140 GCGTTTGGGCTTGCATATCACTAACCAATATCTCAGAGATGTAGAGAGATGCCAG 1139  
 Db 481 GCGTCTGGGGATCGCAATTCATTTAATCTACATCTAGAGATGTGTGAGAGAGATGCCAG 540  
 Qy 1200 AAGGAGAGATATCTTGTCTCAAGATGATTTAGCAGAGGCGCTCTCCAGAGAGA 1239  
 Db 541 AAGGAGAGATCTTACTTGTCTCAAGATGATTTAGCAGAGGCGCTCTATCCGATGAGA 600  
 Qy 1260 CATATTTGCTGGAAGAGTACTGATTAAGTGAAGAGACTTTAT 1301  
 Db 601 TATATTTGCTGGAAGGATGACCGATTAATGAGAAATCTTTAT 642

## RESULT 10

LOCUS AM222027 676 bp mRNA linear EST 18-MAY-2001  
 DEFINITION EST298838 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA  
 clone cLEN6L15, mRNA sequence.

ACCESSION AM222027

VERSION AM222027.1 GI:6533711

KEYWORDS EST.

SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 676)

Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,

Updon,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Renning,C.M.,

Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.

Generation of ESTs from tomato fruit tissue

Unpublished

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES

## SOURCE

1. 676

/organism="Lycopersicon esculentum"

/mol\_type="mRNA"

/cultivar="TA496"

/db\_xref="taxon:4081"

/clone="cLEN6L15"

/tissue\_type="pericarp"

/dev\_stage="red ripe (7-20 days post-breaker)"

/clone\_lib="tomato fruit red ripe, TAMU"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:

XhoI; supplier: Giovannoni; Fruit were tagged at the

breaker stage (first sign of lycopen accumulation on the

blossom end of the fruit) and harvested 7 days

post-breaker (fully red-ripe), 10 days post breaker, and

20 days post-breaker (over-ripe). 20 day fruit which

showed external or internal signs of pathogenesis were

discarded. Fruit were cut in half and the seeds and

locules were discarded prior to freezing the pericarp."

BASE COUNT 214 a 118 c 165 g 179 t

## ORIGIN

Query Match 29.8%; Score 515.8; DB 9; Length 676;

Best Local Similarity 85.5%; Pred. No. 1e-117;

Matches 574; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

Qy 808 CTTGTATGGCCCTTAATGATCCACATTAATCTCGCAAGCTTTAGATAGTGGAGACC 867  
 |||||

Db 2 CTTGTATGAGCCCAACGATCATATATTATCCCGGACCTTAGATAGTGGAGAAAT 61  
 Qy 868 AGGCTGAGAGATATTTTCACTGAGGCGGCCATTTGATATGCTGATGCTTATCCGAT 927  
 Db 62 AGGCTGAGAGATTTTTCATATGCGCGCCATTTGACATGCTCATATGCTTGTCCGAT 121  
 Qy 928 ACTGTCTCAGATTTCTGTGTATTTACGCCATTCAGATATGATTTGAAGAAATCGT 987  
 Db 122 ACAGTTCTTAATCTTCCAGTTGATATTCAGCCATTCAGATATGATTTGAAGAAATCGT 181  
 Qy 988 ATGACCTTGTGAATTCACATACAAACTTTGAGAGAGCTATATCTTATTTACTAT 1047  
 Db 182 ATGACCTTGTGAATTCACATACAAACTTTGAGAGAGCTATATCTTATTTATTTAT 241  
 Qy 1048 GTTGCTGTACTGTAGATGATGATGTTCCAGTTATGAGTATGAGTATGAGCTGATCAAG 1107  
 Db 242 GTTGCTGTACTGTAGATGATGATGTTCCAGTTATGAGTATGAGTATGAGCTGATCAAG 301  
 Qy 1108 GCAACACAGAGAGTATATTAATGCTGCTTTGAGGCTTTCGAAATCAATCAACC 1167  
 Db 302 GCAACACAGAGAGTATATTAATGCTGCTTTGAGGCTTTCGAAATCAATCAATCAACT 361  
 Qy 1168 AATTATCTCAGATGATGAGAGAGATGCCAGAGAGAGAGATATCTTGCCTCAAGAT 1227  
 Db 362 AACTTACTCAGAGATGTTGAGAGAGATGCCAGAGAGAGAGATATCTTGCCTCAAGAT 421  
 Qy 1228 GAATTAGCAGAGAGAGGCTCTCCGACGAGACATATTTCTGTAAGAGTACTGATTAAG 1287  
 Db 422 GAATTAGCAGAGAGAGGCTCTATCCGATGAGATATTTCTGTAAGAGGATGACCGATTA 481  
 Qy 1288 TGGAGAACTTTATGAAGAAACAATTCAGAGGCGGAGAAATCTTTGATGATCAGAG 1347  
 Db 482 TGGAGAACTTTATGAAGAAACAATTCATAGGCAAGAAAGTCTTTGATGAGGCGAGAG 541  
 Qy 1348 AAGAGTCTCAGAACTGAGCTGCTGATGATGATGCTGTGTATACAGGCTGCTGTG 1407  
 Db 542 AAGAGCTCTGAGAAATGAGCTCACTGATGATGATGCTGTGTATGAGGCTGCTGTG 601  
 Qy 1408 TATGCAAGATTTTGAAGAGATTTGAAGCAAGCACTCAACAACCTTCAAGAGAGGCT 1467  
 Db 602 TATGCAAGATTTTGAAGAGATTTGAAGCAAGCAAGCACTTCAACAAGAGAGCA 661  
 Qy 1468 TATGTTAGCAA 1478  
 Db 662 TATGTTAGCAA 672

## RESULT 11

BM412533

LOCUS

DEFINITION EST586860 tomato breaker fruit Lycopersicon esculentum cDNA clone

cLEG60114 5' end, mRNA sequence.

ACCESSION BM412533.1 GI:18264163

VERSION BM412533

KEYWORDS EST.

SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 634)

Alcala,J., Vrebalov,J., White,R., Vison,T., Karameycheva,S.A., Tsai

J., Bouari,O., Kirkness,E., Uteback,T., Van Aken,S., Renning

C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.

Generation of ESTs from tomato fruit tissue, breaker stage (2002)

Unpublished

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

This clone is available through the Clemson University Genomics

Institute



```

Oy 635 AGCCGATATTTGTTTCCAGGAAATTTGGCTTGTGAGTGAACATATGATCGTTTG 694
Db 360 AGCCGATATTTGTTTCCAGGAAATTTGGCTTGTGAGTGAACATATGATCGTTTG 419
Oy 695 GCGAATATGTCAGAGTATGCAAGACATTTTACTTATGAGAACCAAGTATGACCCAG 754
Db 420 GCGAATATGTCAGAGTATGCAAGACATTTTACTTATGAGAACCAAGTATGACCCAG 479
Oy 755 AGAAGAAAGAGCTATCTGGGCAATATATGTTGTGAGAGGAGACGATGAGCTTGTG 814
Db 480 ACAGAGAAAGAGCTATCTGGGCAATATATGTTGTGAGAGGAGACGATGAGCTTGTG 539
Oy 815 ATGGCCCTATGATCCATCCACATTAATCTCCGCAAGCTTTAGTAGTGGAGACAGGCTG 874
Db 540 ATGGCCCTATGATCCATCCACATTAATCTCCGCAAGCTTTAGTAGTGGAGACAGGCTG 599

RESULT 13
BE460889 600 bp mRNA linear EST 18-MAY-2001
LOCUS EST412308 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
DEFINITION clone cLEB36H12, mRNA sequence.
ACCESSION BE460889
VERSION BE460889.1 GI:9505191
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 600)
Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Roming,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
,S.D.
Generation of ESTs from tomato fruit tissue, breaker stage
Unpublished
JOURNAL Contact: CUGI
COMMENT Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
SOURCE location/Qualifiers
1..600
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEB36H12"
/tissue_type="Pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/clone_lib="tomato breaker fruit, TIGR"
/notes="Vector: pBluescriptKmcuadapc; Site 1: EcoRI;
Site 2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopen accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."
BASE COUNT 178 a 99 c 157 g 166 t
ORIGIN
Query Match 28.1%; Score 486.4; DB 10; Length 600;
Best Local Similarity 88.2%; Pred.No. 2.2e-110;
Matches 529; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Oy 650 TTCGAGGGAATTTGGGCTTGTGAGTGAAGCATATGATGTTGGCGAAGATGTGAG 709
Db 1 TTCGAGGGAATTTGGGCTTGTGAGTGAAGCATATGATGTTGGCGAAGATGTGAG 60

Oy 710 AGTATGCAAGACATTTTACTTATGAGAACCAAGCTATGACCCAGAGAGAGAGGCTA 769

```

```

Db 61 AGATGCAAGACGTTTAACTTAGGAATATGCTATGACTCCCGAGAGAAAGGCTA 120
Oy 770 TCTGGCAATATATGTTGTGAGAGAGACGATGATGTTGATGGCCCTATGATCAT 829
Db 121 TCTGGCAATATATGTTGTGAGAGAGACGATGATGTTGATGGCCCTATGATCAT 180
Oy 830 CCACATTAATCTCCGCAAGCTTTAGATATGTTGTGAGAGACGATGATGTTGATCAT 889
Db 181 CATATTAATCTCCGCAAGCTTTAGATATGTTGTGAGAGAGACGATGATGTTGATCAT 240
Oy 890 GCGGCAATGATATATGTTGTGAGAGAGACGATGATGTTGATGGCCCTATGATCAT 949
Db 241 GCGGCAATGATATATGTTGTGAGAGAGACGATGATGTTGATGGCCCTATGATCAT 300
Oy 950 ATATTGACCATTCAGAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 1009
Db 301 ATATTGACCATTCAGAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Oy 1010 ACAAACTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1069
Db 361 ACAAACTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Oy 1070 TGAATGTTCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1129
Db 421 TGAATGTTCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Oy 1130 ATGCTGCTTTGGCTTTAGGCTTTCAGATCAATCAATCAATCAATCAATCAATCAAT 1189
Db 481 ATGCTGCTTTGGCTTTAGGCTTTCAGATCAATCAATCAATCAATCAATCAATCAAT 540
Oy 1190 AAGATGCGAAGAGAGAGAGATATGATGATGATGATGATGATGATGATGATGATGAT 1249
Db 541 AAGATGCGAAGAGAGAGAGATATGATGATGATGATGATGATGATGATGATGATGAT 600

RESULT 14
BE432511 623 bp mRNA linear EST 18-MAY-2001
LOCUS EST399040 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
DEFINITION clone cLEB8H5, mRNA sequence.
ACCESSION BE432511
VERSION BE432511.1 GI:9430354
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 623)
Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Roming,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
,S.D.
Generation of ESTs from tomato fruit tissue, breaker stage
Unpublished
JOURNAL Contact: CUGI
COMMENT Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
SOURCE location/Qualifiers
1..623
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEB8H5"
/tissue_type="Pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/clone_lib="tomato breaker fruit, TIGR"
/notes="Vector: pBluescriptKmcuadapc; Site 1: EcoRI;

```



Site\_2: XhoI; Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

BASE COUNT

191 a 101 c 164 g 167 t

ORIGIN

Query Match 28.0%; Score 483.4; DB 10; Length 623;  
Best Local Similarity 87.2%; Pred. No. 1.2e-109;  
Matches 543; Conservative 0; Mismatches 76; Indels 4; Gaps 1;

QY 674 GTGAAGCATATGATCTTGTGGCGAAGTATGTGCAGATATGCAAAAGACATTTTACTTAG 733  
DB 1 GTGAAGCATATATGTGTGTGGTGAAGTATGTGCAGATATGCAAAAGACATTTTACTTAG 60  
QY 734 GAACCAAGTATATGACCCGAGAGAGAAAGCATATCTGGCCAAATATATGTGTGTGA 793  
DB 61 GAACTATGTATATGATCTCCGAGAGAGAAAGGCTATCTGGCCAAATATATATGTGTGA 120  
QY 794 GGAGAACGATGATGCTTGTGTATGCGCTTATATGATCCCATTAATCTCCGCAAGCTTTAG 853  
DB 121 GAAAGAACATATGATGCTTGTATGCGCCAAAGCATATATATATCCCGGAGACCTTAG 180  
QY 854 ATAGGTGGAGACACGCTGGAAGATATTTTCAATGCGCGCCATTTGATATGCTTGATG 913  
DB 181 ATAGGTGGAGAAATAGGCTGGAAGATGTTTCAATGCGCGCCATTTGATATGCTTGATG 240  
QY 914 CTGCTTATCCGATCTGTCTCCAGATCTCTGTTGATATTTGACCATTCAGAGATATGA 973  
DB 241 GTGCTTTGCGCATATACGTTTCTAATCTTCCAGTTGATATTTGACCATTCAGAGATATGA 300  
QY 974 TTGAAGCATATGATGCTTGTGGAATTCAGATATCAAACTTTCATGAGCTATATC 1033  
DB 301 TTGAAGCATATGCTTGTGATGCTTGTGAGAAATTCAGATATCAAACTTTCAGATATATC 360  
QY 1034 TCTATTTTATC---TATGTTGCTGTACTGTATGATGATGATGCTTTCAGTTATGCT 1089  
DB 361 TTTATTTGTTAATTTATTTATGTTGCTGTGATGCTTGTGATGATGATGCTTTCATTTATGCT 420  
QY 1090 ATTGACCTGATATCAAGGCAACACAGAGATGATATATATGCTTGTGCTTTAGG 1149  
DB 421 ATGCGCCCTGATATCAAGGCAACACAGAGATGATATATATGCTTGTGCTTTAGG 480  
QY 1150 CTTCGAATATCACTAATCAATATATCTCAGAGATGTGAGAGAGATGCCAGAGAGAGA 1209  
DB 481 ATGCAAAATCAATTAATCAATATCTCAGAGATGTGAGAGAGATGCCAGAGAGAGA 540  
QY 1210 GTATACCTGCTCAAGATGATATGACAGGAGGCTCTCCAGACAGACATATTTGCT 1269  
DB 541 GTCTACTGCTCAAGATGATATGACAGGAGGCTCTATCCGATGAATATATTTGCT 600  
QY 1270 GGAAGATGATCTGATATGAGAG 1292  
DB 601 GGGAGGGTGACCGATTAATGAG 623

RESULT 15

AW221932

LOCUS AW221932 685 bp mRNA linear EST 18-MAY-2001  
DEFINITION EST98743 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA  
clone cLENEK7, mRNA sequence.

ACCESSION

AW221932

VERSION AW221932.1 GI:6533616

KEYWORDS

SOURCE

ORGANISM

Lycopersicon esculentum (tomato)

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 685)

Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,

Upton,D., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Romming,C.M.,

Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.

TITLE  
JOURNAL  
COMMENT

Generation of ESTs from tomato fruit tissue  
Unpublished  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

FEATURES

source

1..685

/organism="Lycopersicon esculentum"

/mol\_type="mRNA"

/culivar="TA96"

/db\_xref="taxon:4081"

/clone="cLENEK7"

/tissue\_type="pericarp"

/dev\_stage="red ripe (7-20 days post-breaker)"

/clone\_1b="tomato fruit red ripe, TAMU"

/note="Vector: pBluescript SK(-); Site 1: EcoRI, Site 2: XhoI; supplier: Giovannoni; Fruit were tagged at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit) and harvested 7 days post-breaker (fully red-ripe), 10 days post breaker, and 20 days post-breaker (over-ripe). 20 day fruit which showed external or internal signs of pathogenesis were discarded. Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

BASE COUNT

237 a 113 c 150 g 165 t

ORIGIN

Query Match

Best Local Similarity 27.7%; Score 478.8; DB 9; Length 685;

Matches 537; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 928 ACTGCTCCAGATTTCTGTTGATATTCAGCCATTCAGAGATATGATGAAGAAATGCT 987  
DB 1 ACAGTTTCTAATCTTCCAGTTGATATTCAGCCATTCAGAGATATGATGAAGAAATGCT 60  
QY 988 ATGACCTGTGGAATTCAGATATCAAACTTTCATGAGCTATATCTATTTGTTACTAT 1047  
DB 61 ATGACCTGTGGAATTCAGATATCAAACTTTCATGAGCTATATCTATTTGTTACTAT 120  
QY 1048 GTTGTGTAATCTGTGATATGATGATGCTTCCAGTTATGAGGATATGCACTGAATCAAG 1107  
DB 121 GTTGTGTAATCTGTGATATGATGATGCTTCCAGTTATGAGGATATGCACTGAATCAAG 180  
QY 1108 GCACACAGAGATGATATATATGCTCTTTAGGCTTGAACATCACTAAC 1167  
DB 181 GCACACAGAGATGATATATATGCTCTTTAGGCTTGAACATCACTAAC 240  
QY 1168 AATATATCTCAGAGATGATGAGAAATGCGAAGAGAGAGATATCTTGTCTCAAGT 1227  
DB 241 AACATATCTCAGAGATGATGAGAAATGCGAAGAGAGATATCTTGTCTCTCAAGT 300  
QY 1228 GAATTAAGACAGGAGGCTCTCCGACAGACATATTTGCTGGAAGAGTGAAGTAAAG 1287  
DB 301 GAATTAAGACAGGAGGCTCTATCCGATGAATATTTTGTGGAAGGATGAAGTAAAG 360  
QY 1288 TGAAGAACTTTATGAGAAACAAATTTGAGAGGCGAGAAATTTCTTTGATGACAGAG 1347  
DB 361 TGAAGAACTTTATGAGAAACAAATTTGAGAGGCGAGAAATTTCTTTGATGAGAGAG 420  
QY 1348 AAAGGTGTCAGAACTGGAATCTGTCTGATGATGAGGCTGTGTAAACAGCGCTGTG 1407  
DB 421 AAAGGTGTCAGAACTGGAATCTGTCTGATGATGAGGCTGTGTAAACAGCGCTGTG 480  
QY 1408 TATCGAATATTTGAGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1467  
DB 481 TATCGAATATTTGAGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
QY 1468 TATGTTAGCAAGCAAGAGCTTCTCACTTGGCCATTTGTTATGCAAAATCTCTTGTG 1527  
DB 541 TATGTTAGCAAGCAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 600

Oy 1528 CCCCCTAATAGAACTTCTCTCCACTAGCAAGA 1561  
||| | | | | | | | | | | | | | | |  
Db 601 CTTCTACAAAAAATGCTCTCTCTTCAAGATTA 634

Search completed: January 16, 2004, 04:04:17  
Job time : 3751 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 16, 2004, 04:06:36 ; Search time 5040 Seconds  
(without alignments)  
3571.475 Million cell updates/sec

Title: US-09-847-081b-2  
Perfect score: 2270  
Sequence: 1 MSMSVALLMVVSPITSEVNSG.....IAYAKSLVPNRTSSPLAKT 440

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2888711 seqs, 20454813386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n\_model -DEV=xlh  
-O=/cg2n2.1/USPTO.spool/US09847081/runat.15012004.154336.1853/app\_query.fasta.1.583  
-DB=genemb1 -QFMT=fastap -SUFFIX=rge -MTNMATCH=0.1 -LOOPEL=0 -MODEPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45  
-DOCALLIG=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09847081 @CGN 1.1 3508 @runat.15012004.154336.1853 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-BEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmb1:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_scs:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_rod:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vrt:\*  
38: em\_sy:\*  
39: em\_hngo\_hum:\*  
40: em\_hngo\_mus:\*  
41: em\_hngo\_other:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2192.5	96.6	1573	6	AX657549
2	1827	80.5	1295	8	CAPSYL
3	1751	77.1	1786	8	TOMCBPE
4	1750	77.1	1239	6	A68204
5	1750	77.1	1239	6	AR156064
6	1750	77.1	1239	6	BD005486
7	1744	76.8	1239	6	A68203
8	1744	76.8	1239	6	AR156063
9	1744	76.8	1239	6	BD005485
10	1733	76.3	1614	8	LERIPE
11	1732	76.3	1646	6	A21360
12	1732	76.3	1646	6	AR007503
13	1668	73.5	1355	8	LERYGTOMS
14	1626	71.6	1506	8	AF220218
15	1626	71.6	1689	8	AB037975
16	1606	69.1	1773	8	AF152892
17	1598	70.4	1637	8	HAN308385
18	1595	70.3	2868	6	E15683
19	1590	70.0	1598	8	HAN304825
20	1580.5	69.6	1376	8	AY099482
21	1568.5	69.1	1576	8	CMPSYLMR
22	1568	69.1	1371	8	AP251015
23	1559.5	68.7	1372	6	A48337
24	1557	68.6	1921	6	E15681
25	1544.5	68.0	1853	8	AY056287
26	1542	67.9	1269	6	AX506034
27	1542	67.9	1472	8	BT002084
28	1542	67.9	1703	8	AY085565
29	1542	67.9	1757	8	BT000450
30	1531	67.4	2085	6	E15680
31	1530	67.4	1119	8	TOMPSY2A
32	1526.5	67.2	1769	6	ATHPHYSYN
33	1523	67.1	1932	6	E15682
34	1519	66.9	1548	8	NPPSY
35	1495.5	65.9	4073	8	LESPFS
36	1431.5	63.1	1304	6	AX350858
37	1334.5	58.8	1538	8	AB032797
38	1282.5	56.5	3707	8	LEGTOMS
39	1268.5	55.9	3200	8	AP009954
40	1268.5	55.9	3200	8	AB005238
41	1212	53.4	1060	8	AY024350
42	1064	46.9	681	8	AF196963
43	1059	46.7	686	8	AF158024
44	1054.5	46.5	1278	8	DB091900
45	1052	46.3	1859	8	AP305430

## ALIGNMENTS

RESULT 1



JOURNAL	Biochem. Biophys. Res. Commun. 196 (3), 1414-1421 (1993)
MEDLINE	94071905
PIRMBED	8250898
REFERENCE	2 (bases 1 to 1295)
AUTHORS	Kuntz, M.
TITLE	Direct Submission
JOURNAL	Submitted (30-JUL-1992) M. Kuntz, Inst. de Biologie Molec. des Plantes, 12 rue du Général Zimmer, 67084 Strasbourg, FRANCE
FEATURES	
Source	1. 1295
gene	/organism="Capsicum annuum" /mol_type="mRNA" /strain="lamuyo" /db_xref="taxon:4072" /dev_stage="ripening fruits" 1. 1295 /gene="psy1" 1. 1260 /gene="psy1" /note="subunit structure monomer" /codon_start=1 /product="phycoene synthase" /protein_id="CA48155.1" /db_xref="GI:433994" /db_xref="SWISS-PROT:P37272" /translation="MSVALMWSFCDVSNCTGFLVSVREGNRIFDSSGRNIAQNER IKRGQKQRMWFGSLGAQOQSGRKPSVSAIYATPAGETWSSSEMYDVYLQAA LVKQLSTDELDVAKDIPIEGTGLSEAYDRSEVCAEYAKTFYGTMLMTEREK ALVIAWYCRSTDELDVGPNSHITTPALDWMEDRLDVSGRPFMDLDAISPTVK FPVIOPEPMDIEGMRMDIKRSRYRNDELYLYCYVYAGTYGMSVIMKIAPEKAT TESYNNALAGINQLTNLRDVGDEARRRYVLPDELAQAGLSBDIYAGVTDK WRITMKQIQARFPDEAEKGVTELSAASRPVLASLLYRITLDEIENDVNNFTK RAYVSKPKKLIALPIAAVAKSLVPSRT"
CDS	BASE COUNT 379 a 219 c 361 g 336 t ORIGIN
Alignment Scores:	
Pred. No.:	3.86e-148
Score:	1827.00
Percent Similarity:	89.81%
Best Local Similarity:	83.10%
Query Match:	80.48%
DB:	8 Gaps: 4
US-09-847-081B-2 (1-440) x CAPSY1 (1-1295)	
QY	3 MetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThrGly 22
Db	1 AAGCTGTGTCCTTGTTATGGGTTGTTCTCT--TGTCAGCTCTCAACGGGACAGGA 57
QY	23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42
Db	58 TTCTTGGATCCGTTCCGTGAGGAAACCGGATTTTGAATGCTGC-----CGG 105
QY	43 AspArgAsnLeuMetTrpAsnGlyArgGlyLeuLysLys---GlyGlyArgGlnArgTrpAsn 61
Db	106 CGTAGGAAATTTGGCGTCGCAATGAGAAGATCAACAGACGAGGTGGAAAAAAGGTGAGT 165
QY	62 PheGlySerLeuLeuAlaAspProArgGlySerCysLeuGlyGlySerArgThrGluLys 81
Db	166 TTGGGTCTTAC-----TTGGGAGGAGCAACAACCTGGAAGT 201
QY	82 GlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThrVal 101
Db	202 GGACGGAAATTTTCTGTCGCTTCGTCATCGTGGCTACTCCGGCTGGAGAAATGACATG 261
QY	102 SerSerGluLysLysValTyrAspValValLeuLysGlnAlaAlaLeuValLysArgGln 121
Db	262 TCATCAGAACGAGATGATGATGATGATGATTTTGAAGCAGCAGCCTTGTAAGACAG 321
QY	122 LeuArgSerThrAspAspLeuGluValLysProAspIleValAlaProGlyAsnLeuGly 141
Db	322 CTGAGATCGACCGGATGAGTATGATGAGTAAGAGATATACCTATTCGGGGACCTTTGGGC 381

Oy		142	LeuLeusSerGluAlaTYrAspArgCysGlyValValCysAlaGluTyrAlaIserThrPhe	161
Db		382	TtGTtGAAGTAGACATATGATAGCGTGTAAGTAAGTATGTGCAGAAGTACCCAAAGACGTTT	441
Oy		162	TyrLeuGIYThrLysLeuMetThrProGluAArgArGalaIleTrpAlaIleTyrVal	181
Db		442	TACTTAGGAACGATGCTATAGACTCCGGAGAGAAAGAAAGGCTATCTGGCCATAATACGTA	501
Oy		182	TrpCysArGaGqThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGln	201
Db		502	TGGTCAGAGAAACAGACGAACTGTGTATGGTCCGAATGCATCACTAACATTACTCGGCG	561
Oy		202	AlaLeuAspAcGTTPGLuThrArgLeuGluAspIlePheSerGlyArgProPheAspMet	221
Db		562	GCCTTTAGTAgSTGGAGAGACAGGCTAGAAAGTGTTCCTCAGTGGACGCCCATTTAGCATG	621
Oy		222	LeuAspAlaAlaLeuSerAspThrValSerArGpHeProValAspIleGlnProPheArg	241
Db		622	CTCGATGCTGCTTGTGTCGACACAGTTTCCAATTTCAGTTGATTCAGCCATTACAGA	681
Oy		242	AspMetIleGluGlyMetArgMetAspLeuThrIlePlysSerArgTyrLysThrPheAspGlu	261
Db		682	GATATGATTTAGAGAAATCGTATGACTCTTGAGGAAGTCAAGATACAGAAACCTTGACGA	741
Oy		262	LeuTyrLeuTyrCysTyrTyrValAlaGlyThrValIGlyLeuMetSerValProValMet	281
Db		742	CTATRCCTATATTGTATTACGTTGCTGTAAGGTTGGGTTGATAGTGTTCCAATTATG	801
Oy		282	GlyIleAlaProGluSerLysAlaThrThrglUSeValTyrAspAlaAlaLeuAlaLeu	301
Db		802	GGCATCGCACCTGATCAAAGGCAACAACGGAGACGTAATATATATGCTGCTTTGGCTTGG	861
Oy		302	GlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValIGlyGluAspAlaArgArgGly	321
Db		862	GGGATCGGAATCAGCTGACCACCACTACTTGAAGATGTTGGAAGATGCCCAGAAAGAGA	921
Oy		322	ArgValTyrLeuLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePhe	341
Db		922	AGAGCTCATTTGGCTCAAGATGAATTTACACACAGCAGGCTCTATCCGACGAAGACATATT	981
Oy		342	AlaGlyArGValThrAspLysTrpArgAsnPheMetLysGlnIleGlnArgAlaArg	361
Db		982	GCTGGAAAGATGACCGATTAATATGAGATCTTCATGAAGAAACAATTCAGAGGGCAAGA	1041
Oy		362	LysPhePheAspGluSerGlnLysGlyAlaThrGluLeuAspSerAlaSerArgTrpPro	381
Db		1042	AAGTCTTTTTCAGGGCAGAGAAAAGAGATGACCGAATTGAGGCGACTGATGATGGCCT	1101
Oy		382	ValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGlnAlaAsnAspTyr	401
Db		1102	GTTGTTGGATCTCTGCTTGTGACCGCAGGATCTGCGACGATCGAAACCAATGATCTAC	1161
Oy		402	AsnAsnPheThrArGArGAlaTyrValSerLysProLysLysLeuLeuThrLeuProIle	421
Db		1162	AACAACCTTCACAAAGAGAGCTTATGTGAGCAAAACAAAGAAAGTATTCATTACTATT	1221
Oy		422	AlaTyrAlaLysSerLeuValProProAsnArgThr	433
Db		1222	GCATATGCAAAATCTCTTGTGCTTCTTCAAGAAACA	1257
<b>RESULT 3</b>				
LOCUS	TOMCPBE	1786 bp	mRNA	linear
DEFINITION	Tomato phytoene synthetase mRNA, complete cds.			PLN 27-APR-199
ACCESSION	M84744			
VERSION	M84744.1	GI:170415		
KEYWORDS	carotenoid biosynthesis; phytoene synthetase.			
SOURCE	Lycopersicon esculentum (tomato)			
ORGANISM	Lycopersicon esculentum			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; Jamids; Solanales; Solanaceae; Solanum; Lycopersicon.			

REFERENCE 1 (bases 1 to 1786)  
 AUTHORS Bartley,G.B., Vitanen,P.V., Bacot,K.O. and Scolnik,P.A.  
 TITLE A tomato gene expressed during fruit ripening encodes an enzyme of  
 the carotenoid biosynthesis pathway  
 JOURNAL U. Biol. Chem. 267 (8), 5036-5039 (1992)  
 MEDLINE 92184738  
 PUBMED 1544888  
 COMMENT Original source text: Lycopersicon esculentum (library: Clontech  
 fruit cDNA) breaker fruit stage fruit cDNA to mRNA.  
 FEATURE SOURCE location/Qualifiers  
 1..1786  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4081"  
 /tissue\_type="fruit"  
 /dev\_stage="breaker fruit stage"  
 /cissue\_1fb="Clontech fruit cDNA"  
 314..1552  
 /codon\_start=1  
 /product="phytoene synthetase"  
 /protein\_id="AAA34153.1"  
 /db\_xref="GI:170416"  
 /translation="MSVALLMVSPCDVSNSTFMSVREGNRFPSRRHNVSNR  
 INRGGKQNNNGRFSVSAIATPSERTSQQYDVLRQAALVKGQRTNEL  
 EVKEDIPGNLGLSEAYDRCEGAEYATFNLGTMLTPERRRAIMAIYWCRT  
 DELVDGNASITTPALDRMENRLEDVNGRPFMDLQALSDTNSFPVIOPRDMI  
 EGMMDLRKSYKNPDELILYCYVAGVGVSPINGIAPESKATTESVNNALALG  
 IANQUTNLRDVGEDARGRYRLRODELAQGLSDEDFRGRYTDKRIKQIHNA  
 RPFDEAKGVETLSASRPFVMSLVLYRKILDEIRANDNNFTKAAVYSKSKLLA  
 LPIAVASLVPPTKSLQR"  
 polyA\_site 1746  
 BASE COUNT 594 a 259 c 420 g 513 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2.08e-141 Length: 1786  
 Score: 1751.00 Matches: 344  
 Percent Similarity: 86.67% Conservative: 33  
 Best Local Similarity: 79.08% Mismatches: 34  
 Query Match: 77.14% Indels: 24  
 DB: 8 Gaps: 3  
 US-09-847-081b-2 (1-440) x TOMCPBE (1-1786)  
 QY 1 MetSerMetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGly 20  
 Db 308 CTCGAATGTCGTGTTGCTGTGTATGGGTGTCTCTCT---TGTGACGCTCTCAATGGG 364  
 QY 21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40  
 Db 365 ACAAGTTTCATGGAATCAGTCGGGAGGGAACCGTTTGTGATTGATCG----- 415  
 QY 41 AlaArgAspArgAsnLeuMetTrpAsnGlyValGlyLeuGlyValArgGlnArgTrp 60  
 Db 416 ---AGGCATAGGAATTGGTGTCCATGAGAGAAATCATAGAGGT----- 457  
 QY 61 AsnPheGlySerLeuIleAlaAspProArgTrpSerCysLeuGlyGlySerArgThrGlu 80  
 Db 458 -----GGTGAAGCAAACTAAT 475  
 QY 81 LysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThr 100  
 Db 476 AATGAGCGGAATTTTCTGTACGGCTGCTATTTTGGTACTCATCTGGAAGACGAGC 535  
 QY 101 ValSerSerGluLysValValArgValValLeuLysGlnAlaAlaLeuValLysArg 120  
 Db 536 ATGACATCGAAGACAGATGCTATGATGCTGTTTGAAGCAGCAGCCTTGTGAAGAGC 595  
 QY 121 GlnLeuArgSerThrAspAspLeuGluValLysProAspIleValProGlyAsnLeu 140  
 Db 596 CAACTGAGATCTTACCAATAGATTGAAATGAAGCCGGAATATACCTATTCCGGGAATTTG 655  
 QY 141 GlyLeuLeuSerGluAlaValArgAspArgCysGlyGluValCysAlaGluValArgThr 160

Db 656 GCGTTGTTGAGTAGACATATGATAGTGTCGTAAGTATGTCAGAGTATGCAAGACG 715  
 QY 161 PheTrpLeuGlyThrLysLeuMetThrProGluArgArgValIleLeuPalaIleTyr 180  
 Db 716 TTTAACTTAGGAACATATGCTTAATGACTCCGAGAGAAAGGGCTATCTGGCAATATAT 775  
 QY 181 ValTrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrPro 200  
 Db 776 GTATGCGCAAGAAACAGATGAACTTGTGATGCGCCAAACCATATATATTACCCCG 835  
 QY 201 GlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAsp 220  
 Db 836 GCAAGCTTAAATATGTCGGGAAATAGGCTGAAAGATCTTTCAATAGCGCCCAATTGAC 895  
 QY 221 MetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPhe 240  
 Db 896 ATGCTCATGCTGTGTTGTGTCGATACAGTTCTTAACCTTCAGTTGATTCAGCCATTC 955  
 QY 241 ArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTrpLysThrPheAsp 260  
 Db 956 AGAGATATGATTTGAAGAAATGCGTATGACTTGAGAAATCGAGATTCAAAACCTTGAC 1015  
 QY 261 GlnLeuTrpLeuTrpCysTrpTrpValAlaGlyThrValGlyLeuMetSerValProVal 280  
 Db 1016 GAACATATACCTTATATGTTATATGTTGCTGTGATACGGTGGGTGATGAGTGTCCAAAT 1075  
 QY 281 MetGlyValAlaProGluSerLysValAlaThrThrGluSerValTyrAsnAlaAlaLeuAla 300  
 Db 1076 ATGGATATCCGCCCTGATCAATAAAGGCAACAGAGCGTATATATATATATGCTTGCT 1135  
 QY 301 LeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArg 320  
 Db 1136 CTGGGATATCCCAATTCATTAATTAATCAATCTCAGAGATGTTGGAGAAATGCCAGAGA 1195  
 QY 321 GlyArgValTrpLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIle 340  
 Db 1196 GGAAGAGTCACTTACTGCTCAAGATGAATAGCACAGCAGAGTCACTGATGGAAGATATA 1255  
 QY 341 PheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArgAla 360  
 Db 1256 TTGCTGGAAGGGGTGACCGATTAATGAGAGATCTTTATGAGAAACAAATACATAGGCA 1315  
 QY 361 ArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrp 380  
 Db 1316 AGAAGTCTTTGATATGAGCGAGAAAGCGCTGACAAATTAGAGCTACGTAAGATATTC 1375  
 QY 381 ProValLeuThrAlaLeuLeuLysTrpArgLysIleLeuAspGluIleGluAlaAsnAsp 400  
 Db 1376 CCGTATGAGGACATCTTGTGCTTGACCGCAAAATATAGATGAGATTGAAGCAATGAC 1435  
 QY 401 TyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuPro 420  
 Db 1436 TACAACAACCTTCACAAAGAGACATATGATGACAAATCAAGAAAGTATGATTCATACCT 1495  
 QY 421 IleAlaTyrAlaLysSerLeuValProProAsnArgThrSerSer 435  
 Db 1496 ATTGCATATGCAAAATCTCTTGTCCTCTCAACAAAATCTGCTCT 1540  
 RESULT 4  
 A68204 1239 bp DNA linear PAT 06-MAY-1999  
 LOCUS A68204  
 DEFINITION Sequence 2 from Patent WO9746690.  
 ACCESSION A68204  
 VERSION A68204.1 GI:4759372  
 KEYWORDS  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
 REFERENCE 1 (bases 1 to 1239)  
 AUTHORS Drake,C.R., Bird,C.R. and Schuch,W.W.



```

Db      58  TTTCATGATCAGTCCGGAGGAAACCGTTTTCGATTCATCG-----AGG 105
Oy      43  AspaRgnLeuMetTrrpAnGlyArgIleYslySGlyArgGlnArgTrrpAnphe 62
Db      106  CATAGAAATTTGGTGTCCAAATGAGAAATCAATAGAGGT----- 144
Oy      63  GlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGlyulysGly 82
Db      145  -----GGTGAAGACCAATCAATATATGA 168
Oy      83  SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGlyMetThrValSer 102
Db      169  CGAAATTTTCGTGTCGCTGCTATTTTGGCTACTCCATCTGGAGAACGCGACATGACA 228
Oy      103  SerGlyulysValTyrAspValIleuYslySGlnAlaIleuValIleArgGlnLeu 122
Db      229  TCGGAACAGATGCTATGATGTGTGTGGAGCGACGCTTGTGTGAAGGCGCACTG 288
Oy      123  ArgSerThrAspAspLeuGlyValIlyspProAspIleValProGlyAsnLeuGlyLeu 142
Db      289  AGATTACCAATGATTTAGAACTGAAAGCCGATATACCATTCGCGGAAATTTGGGCTTG 348
Oy      143  LeuSerGlyAlaTyrAspArgCysGlyGlyValCysAlaGlyTyrAlaLysThrPheTyr 162
Db      349  TTGAGTGAAGCATATGATAGTGTGTGTGAGATATGACAGATGCAAGACGTTTAAAC 408
Oy      163  LeuGlyThrIlysLeuMetThrProGlyArgArgAlaIleTrrpAlaIleTyrValTrrp 182
Db      409  TTAGAACTATGCTAATATGATGATCCCGAGAAAGAGGCGCTATCTGGCAATATATGTATGG 468
Oy      183  CysArgGlyThrAspGlyIleuValAspGlyProAsnAlaSerHisIleThrProGlnAla 202
Db      469  TGCAGAAAGACAGATGAACCTTGTGATGGCCAAACGATCATATATATATATATATATAT 528
Oy      203  LeuAspArgTrrpGlyThrArgLeuGlyIleuAspIlePheSerGlyArgProPheAspMetLeu 222
Db      529  TTAGATAGTGGGAAATAGGCTAGAGATGTTTCAATGGGCGGCATTTGACATGCTC 588
Oy      223  AspAlaIleLeuSerAspThrValSerArgPheProValAspIleGlnProPheArgAsp 242
Db      589  GATGGTGTGTGTGCGATACAGTTTCTTAACCTTCCAGTGTATTCACCCATTCGAAAT 648
Oy      243  MetIleGlyIleMetArgMetAspLeuTrrpIlySerArgTyrIlyThrPheAspGlyIleu 262
Db      649  ATGATTTGAAGAAATCGATGACTTGGAAAAATGCAATACAAAACCTTGACGACATCA 708
Oy      263  TrrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGly 282
Db      709  TACCTTATTTGTTATTTATGTTGCTGTCGTCGCTGGGTTGATGAGTTCATATATATGCGT 768
Oy      283  IleAlaProGlySerIlyAlaIleThrArgIleSerValTyrAsnAlaIleuAlaLeuGly 302
Db      769  ATCGCCCTGATCAAGGCAACAGACAGACGCTATATATATGCTGCTTGGCTTGGGG 828
Oy      303  LeuAlaIleGlnLeuThrAsnIleLeuArgAspValGlyGlyAspAlaArgArgIlyArg 322
Db      829  ATCGCAATTCATTTACTTACTTACTCAGAGATGTTGGAGAAAGTCCAGAAAGAGAAAGA 888
Oy      323  ValTrrLeuProGlnAspGlyIleuAlaGlnAlaGlyLeuSerAspGlyAspIlePheAla 342
Db      889  GTCATCTTGCTTCAGATGAAATTAAGACAGAGAGTATCCGATGAGATATATATTTGCT 948
Oy      343  GlyArgValIleThrAspIlyTrrArgAsnMetIlySlyGlnIleGlnArgAlaArgIlyS 362
Db      949  GGAAGGGTGACCGAATTAATGAGATCTTTATGAAGAAACAAATTCATAGGCGCAAGAAAG 1008
Oy      363  PhePheAspGlySerGlyIlySlyGlyValIleThrGlyLeuAspSerAlaSerArgTrrProVal 382
Db      1009  TTTCTTTGATGAGGCGAGAAAGGCGTGCACAGATTTGACTCAGCTAGTAGATTCCTCTGTA 1068
Oy      383  LeuThrAlaLeuLeuLeuTyrArgIlySlyIleLeuAspGlyIleGlyAlaAsnAspTyrAsn 402

```

```

Db      1069  TGGGCATCTTGGTCTTGTAACCGCAAAATACTAGATGATGAGCAATGACTACAC 1128
Oy      403  AsnPheThrArgArgAlaTyrValSerIlyProIlySlyLeuLeuThrLeuProIleAla 422
Db      1129  AACTTCACAAAGAGAGCATATGTGAGCAATTCAAAGAAAGTTGATGATTAATTCATTCGA 1188
Oy      423  TyrAlaIlySerLeuValProProAsnArgThrSerSer 435
Db      1189  TATGCAAAATCTCTTGTGCTCTGCTACAAAACCTGCTCT 1227

RESULT 6
BD005486
LOCUS      BD005486      1239 bp      DNA      linear      PAT 31-JAN-2002
DEFINITION Enhancement of gene expression.
ACCESSION  BD005486
VERSION    BD005486.1 GI:18633857
KEYWORDS   JP 2001501810-A/2.
SOURCE     Lycopersicon esculentum (tomato)
ORGANISM   Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1239)
Drake, C.R., Bird, C.R. and Schuch, W.W.
Enhancement of gene expression
Patent: JP 2001501810-A 2 13-FEB-2001;
ZENBECA LTD

COMMENT
OS   Lycopersicon esculentum (tomato)
PN   JP 2001501810-A/2
PD   13-FEB-2001
PF   23-MAY-1997 JP 1998500302
PR   07-JUN-1996 GB 9611981.3
PI   CAROLINE RACHEL DRAKE, COLIN ROGER BIRD, WOLFGANG WALTER SCHUCH
PC   C12N15/67, C12N15/82, C12N15/29, C07K14/415
CC   Strandedness: Double;
CC   Topology: Linear;
FH   Key
FT   source      1..1239
                     location/Qualifiers
                     /organism='Lycopersicon esculentum (tomato)',
                     /mol_type='genomic DNA'
                     /db_xref='taxon:4081'

BASE COUNT      377 a      207 c      317 g      338 t

ORIGIN
1
...
Alignment Scores:
Pred. No.:      1,58e-141      Length:      1239
Score:          1750.00      Matches:      344
Percent Similarity:      86.84%      Conservative:      32
Best Local Similarity:      79.45%      Mismatches:      33
Query Match:      77.09%      Indels:      24
DB:          6      Gaps:      3

US-09-847-081b-2 (1-440) x BD005486 (1-1239)
Oy      3  MetSerValAlaLeuLeuTrrpValIleSerProThrSerGlyValSerAsnGlyThrGly 22
Db      1  ATGCTGTTCCTTGTATATGAGGTGTTTCCTC--TGTACGCTCTCAATAGGACAAAGT 57
Oy      23  LeuLeuAspSerValArgGlyIlyAsnArgValIlePheValSerSerArgPheLeuAlaArg 42
Db      58  TTTCATGAAATTCATGCTCGGAGGAGAAACGTTTTCATTCATCG-----AGG 105
Oy      43  AspaRgnLeuMetTrrpAnGlyArgIleYslySGlyArgGlnArgTrrpAnphe 62
Db      106  CATAGAAATTTGGTGTCCAAATGAGAAATCAATAGAGGT----- 144
Oy      63  GlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGlyulysGly 82
Db      145  -----GGTGAAGACCAATCAATATATGA 168

```



```

QY      83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThrValSer 102
DB      169 CGAATATTTCTGTAAGCTCTGCTATTTGGCTACTCATCTGAGAAACGCGATGACA 228
QY      103 SerGluValLeuValTyrAspValValLeuLysGlnAlaLeuValLysArgGlnLeu 122
DB      229 TCGGAACAGATGCTTATGATGATGTTTGGAGGACAGCGCTGTTGTAAGGACCACTG 288
QY      123 ArgSerThrAspAspLeuGluValLysProAsp1LeuValProGlyAsnLeuGlyLeu 142
DB      289 AGATCTACCAATAGATAGTAAGTGAAGCCGAGATATACATCTCCGGGCAATTTGGCTTG 348
QY      143 LeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThrPheTyr 162
DB      349 TTGAGTGAAGCATATATAGTGTGTGTAAGTATGTCAGAGATATCAAAAGCGTTAAC 408
QY      163 LeuGlyThrLysLeuMetThrProGluArgArgAlaIleTyrAlaIleTyrValTyr 182
DB      409 TTAGGAACATATGCTAATGAGCTCCGAGAGAAAGGCGTATCTGGCAATATATATGTAG 468
QY      183 CysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHis1IleThrProGlnAla 202
DB      469 TGCAGAGAGACAGATTAACCTGTTGATGCGCCCAACGCATCATATATTACCCCGCAGCC 528
QY      203 LeuAspArgTyrGluThrArgLeuGluAsp1IlePheSerGlyArgProPheAspMetLeu 222
DB      529 TTAGATAGTGGGGAATAATAGGCTAGAAAGATGTTTCAATGGGCGGCATTTGACATGCTC 588
QY      223 AspAlaAlaLeuSerAspThrValSerArgPheProValAsp1LeuProPheArgAsp 242
DB      589 GATGCTGCTTGTCCATACAGTTCATCTTACCTTCCAGTGTGATATTCAGCATTCACAGAT 648
QY      243 Met1LeuGluGlyMetArgMetAspLeuTyrPlySerArgTyrIleThrPheAspGluLeu 262
DB      649 ATGATTGAAGAGATGCTGATGAGCTTGAGAAAATCGAGATACAAAATCTTCAGCAACTA 708
QY      263 TyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGly 282
DB      709 TACCTTATTTGTTATATATGTTGCTGCTAGCGTTGGTGTGATGATGATTCATATATGGGT 768
QY      283 IleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGly 302
DB      769 ATCGCCCCGGAATCAAAAGGCAACACAGAGCGTATATATGCTCTTGGCTCTGGCG 828
QY      303 LeuAlaAsnGluLeuThrAsn1IleLeuArgAspValGlyGluAspAlaArgArgGlyArg 322
DB      829 ATCGCAATCAATTAATTAACATCTCAAGATGTTGAGAGATGCCAGAAAGGAGAGA 888
QY      323 ValTyrLeuProGluAspGluLeuAlaGlnAlaGlyLeuSerAspGluAsp1IlePheAla 342
DB      889 GTCTACTTCCCTCAATGATGATTAAGCAACGCGATCTATCCGATGAAATATATTTGCT 948
QY      343 GlyArgValThrAspLysTyrPheAsnPheMetLysLysGln1LeuAlaArgAlaArgLys 362
DB      949 GGAAGGCTGACGATTAATGAGAAATCTTATGAAAGAAACAATAATCATGAGGCAAGAAAG 1008
QY      363 PhePheAspGluSerLysGlyValThrGluLeuAspSerAlaSerArgTyrProVal 382
DB      1009 TTCTTTGATGAGCGACAGAAAGCGTGAAGAAATTGAGCTCACTGATGATGATCTCTGTA 1068
QY      383 LeuThrAlaLeuLeuLeuTyrArgLys1IleLeuAspGlu1LeuAlaAsnAspTyrAsn 402
DB      1069 TGGGCACTTTGCTCTGTAACCGCAAAATATGATGATGATGATGATGATGATGATGATGAT 1128
QY      403 AsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIleAla 422
DB      1129 AACTTCACAAAGACATATATGAGCAAAATCAAAAGATGATGATGATGATGATGATGATGATGAT 1188
QY      423 TyrAlaLysSerLeuValProProAsnArgThrSerSer 435
DB      1189 TATGCAAAATCTTGTGCTCTCTACAAAATGCTCTCT 1227

```

RESULT 7

```

A68203
LOCUS      A68203                1239 bp      DNA      linear      PAT 06-MAY-1999
DEFINITION Sequence 1 from Patent WO9746690.
ACCESSION  A68203
VERSION     A68203.1  GI:4759371
KEYWORDS
SOURCE      unidentified
            unidentified
ORGANISM    unidentified
REFERENCE   1 (bases 1 to 1239)
AUTHORS     Drake, C.R., Bird, C.R. and Schuch, W.W.
TITLES      ENHANCEMENT OF GENE EXPRESSION
JOURNAL     Patent: WO 9746690-A 1 11-DEC-1997;
            ZENECA LTD (GB)
FEATURES
source      1..1239
            /organism="unidentified"
            /mol_type="genomic DNA"
            /isolate="SYNTHETIC DNA"
            /db_xref="taxon:32644"
BASE COUNT  343 a      247 c      320 g      329 t
ORIGIN
Alignment Scores:
Pred. No.:      5,19e-141      Length:      1239
Score:          1744.00      Matches:      343
Percent Similarity: 86.61%      Conservative: 32
Best Local Similarity: 79.21%      Mismatches: 34
Query Match:    76.83%      Indels:      24
DB:              6      Gaps:      3
US-09-847-081B-2 (1-440) x A68203 (1-1239)
QY      3 MetSerValAlaLeuLeuTyrValLysSerProThrSerGluValSerAsnGlyThrGly 22
DB      1 ATGACCGTGGCACCTTCTTGGTGTGAGCCCA--TCCGATGATAGTAACCGCACTTCA 57
QY      23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42
DB      58 TTTATGAGAGGTGTAGAGAAAGGTATATGATCTTCGACAGTTCT-----CGT 105
QY      43 AspArgAsnLeuMetTyrAsnGlyArg1IleLysLysGlyGlyArgGlnArgTyrAsnPhe 62
DB      106 CACCGTAACCTTGATGTAACGAACGTAATAACAGGGG----- 144
QY      63 GlySerLeu1IleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluLysGly 82
DB      145 -----GGAGGTAAACAGACAAACACCGGT 168
QY      83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThrValSer 102
DB      169 AGAAAGTCTCAGTGAATCGCAATCCTTGAACACCTAGGCGGTGAGAGAACTATGACT 228
QY      103 SerGluLysValTyrAspValValLeuLysGlnAlaLeuValLysArgGlnLeu 122
DB      229 AGCGAGCAAAATGCTTACGACGCTCTACTCTGCAAGTGCACATGTTAAAGTCACTTA 288
QY      123 ArgSerThrAspAspLeuGluValLysProAsp1IleValProGlyAsnLeuGlyLeu 142
DB      289 CTTACTACTCAACGACTTGAAGGTTAACTGACATTCCAATACCTGAAACCTTGACTT 348
QY      143 LeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThrPheTyr 162
DB      349 TTCTCTGAGGCTTACGACAGATGCGGAGAGGTTTGCGAGAAATGAGCTTAAACCTTCAAT 408
QY      163 LeuGlyThrLysLeuMetThrProGluArgArgAlaIleTyrAlaIleTyrValTyr 182
DB      409 TTGGGTACACATGTTATATGACACCAAGAGGCGTGTGCAATATGCGCTATTTAGCTTGG 468
QY      183 CysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHis1IleThrProGlnAla 202
DB      469 TGTAGCGTACTGACGAGATGATGAGCGGACCTAATGCTAGTACATACACACCCGCTGCT 528

```

Qy 203 LeuaspargtrpGluThrArgLeuGluaspIlePheSerGlyArgProPheAspMetLeu 222  
 Db 529 CTGACAGATGGAGAACCGTTTGGAGAGCGTTTAACGGACGCTTTCGATATGTTG 588  
 Qy 223 AspaIaIaleuSerAspThrValSerArgPheProValAspIleGlnProPheArgasp 242  
 Db 589 GACGAGACCTTAGACACCTGTGACAAATTCCTGTGACATCCAACTTTTCGGGAC 648  
 Qy 243 MetIleGluGlyMetArgMetAspLeuThrPlySerArgGlyTyrIlePheAspGluLeu 262  
 Db 649 ATGATCGAGGGCATAGAAATGATCTTCGTAAATCTCGTATTAAGAAATTTGATGATG 708  
 Qy 263 TyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGly 282  
 Db 709 TATTTGATCTGCTACTACGTGGACGAAACCGTGGCCCTTATAGTCATGCTCATGGGA 768  
 Qy 283 IleAlaProGluSerIleValAlaThrThrGluSerValIleAsnAlaIleuAlaLeuGly 302  
 Db 769 ATTGCACAGAGAGTAAGCTACTGATCTGTTTACCGCAGCATAGCATTAAGCT 828  
 Qy 303 LeuAlaAsnGluLeuThrAsnIleLeuArgAspValGlyIleAspAlaArgArgGlyArg 322  
 Db 829 ATAGCTAACCACTTACAAATATCTTGAGGAGCGTGGGTGAGACGCACTTGGGTCGT 888  
 Qy 323 ValTyrLeuProGluAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAla 342  
 Db 889 GTGATCTCCACAGACGAGCTGCTCAAGCTGATTTGATGACGAGACATTTTGGCA 948  
 Qy 343 GlyArgValThrAspPlyThrArgAsnPheMetIleValGlnIleGlnArgAlaArgIys 362  
 Db 949 GGTCTGTGTACAGACAGATGGAGATTTTCATGAAAAAGCAATTCACCGTCTGTAAA 1008  
 Qy 363 PhePheAspGluSerGluIleGlyValThrGluLeuAspSerAlaSerArgTyrProVal 382  
 Db 1009 TTTTTCGACGAAGCTGAAAAGGAGCTTACTAGCTTTCTAGTCATCAAGTTTCAATT 1068  
 Qy 383 LeuThrAlaLeuLeuLeuTyrArgIysIleLeuAspGluIleGluAlaAsnAspTyrAsn 402  
 Db 1069 TGGGCGACCTTGTCTCTATAGAAAGATTTTGGACGAATTCGAGCTTATATAT 1128  
 Qy 403 AsnPheThrArgArgAlaTyrValSerIleProIleValLeuThrLeuProIleAla 422  
 Db 1129 AATTTTACTAAACGCTTACGTTCTTAAGAGCAAAAAAATTATGCTTCCATTCCT 1188  
 Qy 423 TyrAlaIysSerIleuValProProAsnArgThrSerSer 435  
 Db 1189 TACGCTAAGAGCTTGTTCCACCACTAAGACGCTAGC 1227  
 RESULT 8  
 ARI56063 1239 bp DNA linear PAT 08-AUG-2001  
 LOCUS AR156063  
 DEFINITION Sequence 1 from patent US 6239331.  
 ACCESSION AR156063  
 VERSION AR156063.1 GI:15124116  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1239)  
 AUTHORS Drake, C.Rachel, W. Bird, C.Roger, and Schuch, W.Walter.  
 TITLE Enhancement of tomato phytoene synthase gene expression with a modified DNA  
 JOURNAL Patent: US 6239331-A 1 29-MAY-2001;  
 FEATURES Location/Qualifiers  
 SOURCE 1..1239  
 /organism="unknown"  
 BASE COUNT 343 a 247 c 320 g 329 t  
 ORIGIN  
 Alignment Scores: 5.19e-141 Length: 1239  
 Pred. No.: 1744.00 Matches: 343  
 Percent Similarity: 86.61% Conservative: 32

Best Local Similarity: 79.21% Mismatches: 34  
 Query Match: 76.83% Indels: 24  
 DB: 6 Gaps: 3  
 US-09-847-081b-2 (1-440) x ARI56063 (1-1239)  
 Qy 3 MetSerValAlaLeuLeuThrValValSerProThrSerGluValSerAsnGlyThrGly 22  
 Db 1 ATGAGCTGGGCACTTCTTGGTGTGAGCCCA---TGCATGTGAATTAACGGACATTCA 57  
 Qy 23 LeuLeuAspSerValaGluGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42  
 Db 58 TTTATGAGAGGTGTGAGAGAGGTAATGATTTCTTCAAGTTCT-----CGT 105  
 Qy 43 AspArgAsnLeuMetTPanGlyArgIleIleIleValGlyArgGlnArgTrpAsnPhe 62  
 Db 106 CACCGTAACTCTGTAGTAAAGAACTGATTAACAGGGGA----- 144  
 Qy 63 GlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluIleGly 82  
 Db 145 -----GGAGGTAAACAGAACAAACGCT 168  
 Qy 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThrValSer 102  
 Db 169 AGAAAGTTCTCAGTTAGATGACGAATCTTGCAACACCTTGCGGTGAGAGAACTATGACT 228  
 Qy 103 SerGluIleValValTyrAspValValLeuIleValAlaIleuValIleValArgGlnLeu 122  
 Db 229 AGCAGCAAAATGTTGTCAGACGTGTCCTTCTGTAACCTGACCTAGTTAAACGTCACTTA 288  
 Qy 123 ArgSerThrAspLeuGluValIleValProAspIleValValProGluAsnLeuGlyLeu 142  
 Db 289 CGTAGTACTAACGAACCTTGAGGTTAACTGACCTGCAATCTGAAACCTTGACTT 348  
 Qy 143 LeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGlyTyrAlaIleThrPheTyr 162  
 Db 349 CTTTCTAGGCTTACGACAAATGCGAGAGTTTGGCCAGAAATACGCTTAAACCTTTCAAT 408  
 Qy 163 LeuGlyThrIleValMetThrProGluArgArgAlaIleThrPalaIleTyrValTyr 182  
 Db 409 TTGGGTCCATGTTGATGACCAAGAAAGCGCTGTCCAATATGCGCTATTTACGTTGG 468  
 Qy 183 CysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGlnAla 202  
 Db 469 TGTAGGGGTACTGACGAGTTATGTGACGACCTTAATCTGATTAACATTAACCCGCTGCT 528  
 Qy 203 LeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeu 222  
 Db 529 CTTCACAGATGGAGAACCGTTTGGAGACGTGTTTAAACGACACCTTTCGATATGTTG 588  
 Qy 223 AspaIaIaleuSerAspThrValSerArgPheProValAspIleGlnProPheArgasp 242  
 Db 589 GACGAGACCTTAGACACCTGTGACAAATTCCTGTGACATCCAACTTTTCGGGAC 648  
 Qy 243 MetIleGluGlyMetArgMetAspLeuThrPlySerArgGlyTyrIlePheAspGluLeu 262  
 Db 649 ATGATCGAGGGCATAGAAATGATCTTCGTAAATCTCGTATTAAGAAATTTGATGATG 708  
 Qy 263 TyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGly 282  
 Db 709 TATTTGATCTGCTACTACGTGGACGAAACCGTGGCCCTTATAGTCATGCTCATGGGA 768  
 Qy 283 IleAlaProGluSerIleValAlaThrThrGluSerValIleAsnAlaIleuAlaLeuGly 302  
 Db 769 ATTGCACAGAGAGTAAGCTACTGATCTGTTTACCGCAGCATAGCATTAAGCT 828  
 Qy 303 LeuAlaAsnGluLeuThrAsnIleLeuArgAspValGlyIleAspAlaArgArgGlyArg 322  
 Db 829 ATAGCTAACCACTTACAAATATCTTGAGGAGCGTGGGTGAGACGCACTTGGGTCGT 888  
 Qy 323 ValTyrLeuProGluAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAla 342  
 Db 889 GTGATCTCCACAGACGAGCTGCTCAAGCTGATTTGATGACGAGACATTTTGGCA 948

QY 363 GYAAGVAlThrAspLysTrpArgAsnMetLysLysGlnIleGlnArgAlaArgLys 362  
 Db 949 GGTGGTGTACAGACAGAGTGGAGGATTTTCATGAAAAACAGATTACCGTCTGTATA 1008  
 QY 363 PhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpProVal 382  
 Db 1009 TTTTTCGACGAAGCTGAAAAAGGAGTTACTGAGCTTTCTAGTCATCAAGGTTCCAGTT 1068  
 QY 383 LeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGlnAlaAsnAspTyrAsn 402  
 Db 1069 TGGGCGACGCTGTGTCTATAGAAAAGATTGGACGAATCGAGGCTTAACATATATAT 1128  
 QY 403 AsnPheThrArgArgAlaTyrValSerLysProLysLeuLeuThrLeuProIleAla 422  
 Db 1129 AATTTTACAAACGCTCTTACGTTCTTCAAGACAAAACTTATCGCTCTTCCATCGCT 1188  
 QY 423 TyrAlaLysSerLeuValProProAsnArgTrpSerSer 435  
 Db 1189 TACGCTAAGAGCTGTGTTCCACCACTAAGACAGCTAGC 1227

RESULT 9  
 LOCUS BD005485 1239 bp DNA linear PART 31-JAN-2002  
 DEFINITION Enhancement of gene expression.  
 ACCESSION BD005485  
 VERSION BD005485.1 GI:18633856  
 KEYWORDS JP 2001501810-A/1.  
 SOURCE unclassified  
 ORGANISM unclassified  
 1 (bases 1 to 1239)  
 REFERENCE Drake, C.R., Bird, C.R. and Schuch, W.W.  
 AUTHORS Enhancement of gene expression  
 TITLE Patent: JP 2001501810-A 1 13-FEB-2001;  
 JOURNAL ZENECA LTD  
 COMMENT OS Unidentified  
 PN JP 2001501810-A/1  
 PD 13-FEB-2001  
 PE 23-MAY-1997 JP 1998500302  
 PR 07-JUN-1996 GB 9611981.3  
 PI CAROLINE RACHEL DRAKE, COLIN ROGER BIRD, WOLFGANG WALTER SCHUCH  
 CC C12N15/67, C12N15/82, C12N15/29, C07K14/415  
 CC Strandness: Double;  
 CC Topology: Linear;  
 FH Key Location/Qualifiers  
 FT source 1..1239  
 FT Location/Qualifiers  
 1..1239  
 /organism="unclassified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"

BASE COUNT 343 a 247 c 320 g 329 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 5.19e-141 Length: 1239  
 Score: 1744.00 Matches: 343  
 Percent Similarity: 86.61% Conservative: 32  
 Best Local Similarity: 79.21% Mismatches: 34  
 Query Match: 76.83% Indels: 24  
 DB: 6 Gaps: 3

US-09-847-081b-2 (1-440) x BD005485 (1-1239)

QY 3 MetSerValAlaLeuLeuTyrValValSerProThrSerGluValSerAsnGlyThrGly 22  
 Db 1 ATGAGCGTGGCACTTCTTTGGGTGGTGAACCCCA---TGGGATGTAGTAAACGGCACTTCA 57  
 QY 23 LeuLeuAspSerValArgGluLysAsnArgValPheValSerSerArgPheLeuAlaArg 42  
 Db 58 TTTATGCGAGTGTGTGAGAGAGTATATAGATTCTTCGACAGTTCT-----CGT 105

QY 43 AspArgAsnLeuMetTrpAsnGlyValGlyLysGlyArgGlnArgTrpAsnPhe 62  
 Db 106 CACCGTAACTCTGTAGTAAACAGACGTATTAACAGGGA----- 144  
 QY 63 GlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluLysGly 82  
 Db 145 -----GGAGGTAAACAGCAACAACCGT 168  
 QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyLysMetThrValSer 102  
 Db 169 AGAAAGTCTCAGTTAGTACAGCAATCTTCGACACCACTAGCGGTGAGAGAACTATGACT 228  
 QY 103 SerGluLysValTyrAspValValLeuLysGlnAlaIleLeuValLysArgGlnLeu 122  
 Db 229 ACCGCAAGAAAGGTGTACGACGCTGCTACTTGTCAAGTCGACATGTTAAACGTACGTTA 288  
 QY 123 ArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeuGlyLeu 142  
 Db 289 CGTACTACTTAACGAACCTTGAAGTTAACTGACATTCGAAATCCTGAAACCTTGACTT 348  
 QY 143 LeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThrPheTyr 162  
 Db 349 CTTTCTGAGGCTTACGACAGATGCGGAGAGGTTTGGCAGAAATACGTTAAACCTTCAAT 408  
 QY 163 LeuGlyThrLysLeuMetThrProGluArgArgArgAlaIleTyrAlaIleTyrValTrp 182  
 Db 409 TTGGGTACCATGTTGATGATACACGAAAGGGGTGCGCAATATGGCTATTTTACCTTTGG 468  
 QY 183 CysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGlnAla 202  
 Db 469 TGTAGCGCTACGACAGACTTGTAGTGGACGCACTTAATGTTATCATTAACCCGCTGCT 528  
 QY 203 LeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeu 222  
 Db 529 CTTGACAGATGGGAGAACCGTTTGGAGCAGCGTTTAAACGCGACAGACTTTCGATATGTTG 588  
 QY 223 AspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArgAsp 242  
 Db 589 GACGAGACACTTAGACACTGTGAGCAATTTCCCTGTGACATTCACCTTTTGGGAC 648  
 QY 243 MetIleGluGlyMetArgMetAspLeuThrPheSerArgTyrLysThrPheAspGluLeu 262  
 Db 649 ATGATCGAGGGGATAGATGATGCTTGTAACTGCTTATTAAGAAATTTTGAATGATG 708  
 QY 263 TyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGly 282  
 Db 709 TATTTGTACTGCTACTACGTGGCAGGAACCGTGGGCTTATGTCACTGCTATCATGGA 768  
 QY 283 IleAlaProGluSerLeuValAlaThrThrGluSerValTyrAsnAlaIleLeuAlaLeuGly 302  
 Db 769 ATTGACACAGAGAGTAAAGCTACTACTGAATCTGTTTACACCGCAGCACTGACATTAAGT 828  
 QY 303 LeuAlaGlnLeuLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgGlyArg 322  
 Db 829 ATAGCTAACAGCTTACAAATATCTTGAGGAGACGTGGGTAGAGCGCACGTAAGGGTCT 888  
 QY 323 ValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAla 342  
 Db 889 GTGTATCTCCACAGAGAGAGACTCGCTCAACGTGATGAGTGAAGACGACGACATTTTCCCA 948  
 QY 343 GlyArgValThrAspLysTrpArgAsnMetLysLysGlnIleGlnArgAlaArgLys 362  
 Db 949 GGTGTGTATCAAGACAGTGAAGATTTTCAATGAAAAAGCAGATTACCGTGTCTGTAA 1008  
 QY 363 PhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpProVal 382  
 Db 1009 TTTTTCGACGAAGCTGAAAAAGGAGTTACTGAGCTTTCTAGTCATCAAGGTTCCAGTT 1068  
 QY 383 LeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGlnAlaAsnAspTyrAsn 402  
 Db 1069 TGGGCGACGCTGTGTCTATAGAAAAGATTGGACGAATCGAGGCTTAACATATATAT 1128

QY	403	Aspethtraxarxat1TyValSerlyspolylytleuThleuProtlea	422
DB	1129	AAATTACTAAAGCTTACCTTTCTAAGACCAAAAACCTATGCTTCTCAATGCT	1188
QY	423	TyrAlaTySerIueValProProAnaRgThrSerSer	435
DB	1189	TACGCTAAGAGCTTGCTTCCACCACTAAGACAGCTAGC	1227
RESULT 10			
LOCUS	LERIPE	1614 bp	linear
DEFINITION	Tomato fruit ripening specific mRNA.		PLN 11-MAY-1995
ACCESSION	Y00521		
VERSION	Y00521.1		
KEYWORDS	GI:19340		
SOURCE	unidentified reading frame.		
ORGANISM	Lycopersicon esculentum (tomato)		
REFERENCE	Lycopersicon esculentum		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.		
TITLE	1 Ray, J., Bird, C., Maunders, M., Grierson, D. and Schuch, W.		
JOURNAL	Sequence of PTOM5, a ripening related cDNA from tomato		
MEDLINE	Nucleic Acids Res. 15 (24), 10587 (1987)		
PUBMED	88096591		
REFERENCE	3697097		
AUTHORS	2 (bases 1 to 1614)		
TITLE	Bird, C.R.		
JOURNAL	Direct Submission		
DEFINITION	Submitted (09-DEC-1987) Bird C.R., Imperial Chemical Industries PLC, Plant biotechnology group, Po Box 11, The Heath, Runcorn, Cheshire, WAT 40E, U.K		
FEATURES			
source	Location/Qualifiers		
	1..1614		
	/organism="Lycopersicon esculentum"		
	/mol_type="mRNA"		
	/strain="Mill"		
	/cultivar="Alisa Craig"		
	/db_xref="taxon:4081"		
	/clone="PTOM5"		
	201..1439		
	/note="tomato fruit ripening specific URF"		
	/codon_start=1		
	/protein_id="CAA68575.1"		
	/db_xref="GI:19341"		
	/db_xref="SWISS-PROT:P08196"		
	/translation="MSVALIHWSPCDVSNGTSPMESVREGNRPFDSSRRHLVSNERRINRGKQKNNGRKFSVAIIATPSSGRITSSBQMYDVVLRQALVKROLRSTNEIEVKPDIPIPGNIGLSEAYDRGCEVCACTAFNIGTLMTPERRAIIATYWCRRATDELVGPAASTYTPPALRWNRRLBDEVNGRPDMLDGLSDTVSNFEPVDIOPRDMDEGMWDLRKSRYNKDFLXICVYAAAGVGLMSVINGIABESKATTESVYNALAIACIGANQTLIRLDVGEDARGRVYILPDDELAQGLSDESDIPAGRVTDKRIIPKKQIHRARKFEAEKGVTELSASRFPVMAVLVLRKILDEIANDYNNFTKRAYVSKSRQVDCIYYICICKISCASYKNAASLQR"		
	246..254		
	/note="putative glycosylation site"		
	702..710		
	/note="putative glycosylation site"		
	1329..1337		
	/note="putative glycosylation site"		
	1419..1427		
	/note="putative glycosylation site"		
BASE COUNT	515 a 241 c 388 g 470 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	6.5e-140	Length:	1614
Score:	1733.00	Matches:	346
Percent Similarity:	85.68%	Conservative:	31
Best Local Similarity:	78.64%	Mismatches:	36
Query Match:	76.34%	Indels:	27
DB:	8	Gaps:	4

US-03-847-081B-2 (1-440) x LER1PE (1-1614)

QY	1	MetserNetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAngly	20
Db	195	CTCGAAATGTCGTGGTCCCTGTGTATAGGGGTGTTCTCCT---TGTAGGCTCTCAATATGGG	251
QY	21	ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu	40
Db	252	ACAAATTTCATGAGATCACTCCGGAGGAGAAACCTTTTGTGATTCAATCG-----	302
QY	41	AlaArgAspArgAsnLeuMetTrpAsnGlyArgGlyLeuValGlyValArgGluArgTrp	60
Db	303	---AGGCATAGCAATTTGCTGCCATAGAGAAATCAATAGAGCT-----	344
QY	61	AsnPheGlySerLeuLeuAlaAspProArgTyrSerCysLeuGlyGlySerArgTrpGlu	80
Db	345	-----GGTGAAGAACCAACTAAT---:::	362
QY	81	LyseGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyLeuMetThr	100
Db	363	AATGACCGAAATTTTCTGACGGCTGCTCTATTGTCCTACTCCATCTGAGAACGAGAG	422
QY	101	ValSerSerGluValValValTyrAspValValLeuValSGlnAlaAlaLeuValLysArg	120
Db	423	ATGACATCGAAGACAGATGCTCTATCATGTGTTTGAAGCAGACGCTTGTTGTAAGAG	482
QY	121	GlnLeuArgSerThrAspAspLeuGluValLysProAspIleValValProGlyValLeu	140
Db	483	CACTGAGATCTCAATGAGTTGTAAGATGAAAGCCGATATATCTAATTCGGGGAATTTG	542
QY	141	GlyLeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThr	160
Db	543	GCGTTGTGAATGAAAGCATATGATAGGTGTGTGAAGTATGACAGATATCCAAAGAG	602
QY	161	PheTyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrpAlaIleTyr	180
Db	603	TTTAACTTAGAATCTATCTATATGACTCCGAGAGAAAGGCTATCTGGGCAATATAT	662
QY	181	ValTyrCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrPro	200
Db	663	GTAAGTGCAGAAAGACAGATGAATCTTGTATGGCCCAACGCATCATATTTACCCG	722
QY	201	GlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyValArgProPheAsp	220
Db	723	GCAGCTTAGATAGGTGGGAAATAGCTAGACAGATGTTTCAATCGCGGCATTTGAC	782
QY	221	MetLeuAspAlaAlaLeuSerAspThrAlaSerArgPheProValAspIleGlnProPhe	240
Db	783	ATGCTCGAGTGTCTTGTCCGATACATGTTCTTAATTTCCAGTTGATATTCAGCATTC	842
QY	241	ArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysThrPheAsp	260
Db	843	AGAGATATGATTGAAGAGATCGTATGACCTTGAGAAAAATCAGATACAAAACCTTCGAC	902
QY	261	GluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProVal	280
Db	903	GAACTATACCTTTATTTATTTATTTATTTCTGTGATGGTGGTGTATGATGATTTCCAAT	962
QY	281	MetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAla	300
Db	963	ATGGGTATCGCCCTGATCAAGGCAACAACAGAGACGTATATTAATGCTGCTTTGGCT	1022
QY	301	LeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArg	320
Db	1023	CTGGGGAATCGCAATCAATTAATCAATCACTCAAGATGTTGGAGAGATATCCAGAGAA	1082
QY	321	GlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIle	340
Db	1083	GGAGAGCTATCTTGCTCCCAAGATGAATTAAGACAGGACGCTATTCGATGAAGATATA	1142
QY	341	PheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysLeuGlnIleGluArgAla	360
Db	1143	TTTGTCTGAAAGGATGACGATTAATATGAGAAATCTTTATGAAGAAACAATATCATATGGCA	1202

```

QY      361 ArgLysPhePheAspGluSerGluValThrGluLeuAspSerAlaSerArgTTP 380
DB      1203 AGAAGGTTCTTTCATGAGCAGAGGAGGATGACGAAATGAGCTCAGCTAGTATTC 1262
QY      381 ProValLeuThrAlaLeuLeuLeuValArgLysIleLeuAspGluIleGluAlaAsnAP 400
DB      1263 CCTGATGGGCACTTGGCTTGTGACCGCAAAATCTAGATGAGATTGACCAATGAC 1322
QY      401 TyrAsnAsnPhenThrArgArgAlaTyrValSerLeuProLys-LysLeuLeuThrLeuP 420
DB      1323 TACACAACTTCACCAAGAGAGCATATGAGGCAAAATCAAGCAAGTGAATTCATTACC 1382
QY      420 oilealaTyrAlaLysSerLeuValProProAsnArgThrSerSerProLeuAlaLys 439
DB      1383 TATTCATATGCAAAATCTCTTGTGCTCTCT-----ACAAAATGCTCTCTTCAA 1434

RESULT 11
LOCUS      A21360      1646 bp      mRNA      linear      PAT 06-JUN-1994
DEFINITION L.esculentum pTOM5 mRNA.
ACCESSION  A21360
VERSION     A21360.1 GI:512395
KEYWORDS
SOURCE      Lycopersicon esculentum (tomato)
ORGANISM    Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1646)

REFERENCE   1
  AUTHORS   DNA, CONSTRUCTS, CELLS AND PLANTS DERIVED THEREFROM
  TITLE     Patent: WO 9109128-A 1 27-JUN-1991.
  JOURNAL   Location/Qualifiers
            1. .1646
              /organism="Lycopersicon esculentum"
              /mol_type="mRNA"
              /variety="Alisa Craig"
              /db_xref="taxon:4081"
            1. .1646
              /gene="pTOM5"
              201. .1439
              /note="involved in the carotenoid pathway"
              /codon_start=1
              /protein_id="CAA01548.1"
              /db_xref="GI:512396"
              /translation="MSVALLMVSPCDVNGTSMFESVREGNRFDSRHRLVSENER
              IRRGGKOTNNGRKFSVSAIILATPSGEKRTMSQWVYDVVLRQALVROLSTNEL
              EVKPIPIPGNIGLSEAYDRGECVCAEAKTPNIGTMPEERRALIAIYVRCRT
              DELVNDLRKSRKNDDELTYCYIYAGIVGLMSVPIWGLAPSKATTESVYAAALMG
              IANQLTNIIRDVGEADRGRVYLPQDLAQGLSDEIDFAGRTDKMRIFMKQIHR
              RKFPEAEKGVTELSASRFPYMASIVLYRKILDEIDANNYNTKRAVSVSKSQVDC
              IYCTCKISCAASYKTASLQR"
            1630
              /gene="pTOM5"

BASE COUNT  529 a      249 c      388 g      480 t
ORIGIN
polyA_site
/gene="pTOM5"

Alignment Scores:
Pred. No.:      8, 13e-140      length:      1646
Score:          1732.00      Matches:      346
Percent Similarity: 85.68%      Conservative: 31
Best Local Similarity: 78.64%      Mismatches: 36
Query Match:     76.30%      Indels:      27
DB:              6      Gaps:      4

US-09-847-081B-2 (1-440) x A21360 (1-1646)
QY      1 MetSerMetSerValAlaLeuLeuTTPValValSerProThrSerGluValSerAnGly 20
      :::|

```

```

DB      195 CTCAGAAATGTCCTGTTGCTTTATGGGTTGTTCTCTCT---TGTAGCGTCAAAATGGG 251
QY      21 ThrGlyLeuLeuAspSerValArgLysIleValAsnArgValPheValSerSerArgPheLeu 40
DB      252 ACAAGTTTCATGAATCAAGTCCGGAGGGGAAACCGTTTTTGTGATTCACG----- 302
QY      41 AlaArgAspArgAsnLeuMetCysPheArgLysIleValLysGlyGlyArgGlnArgTTP 60
DB      303 ---AGGCATAGGAATTTGGTGTCCATAGAGAAATCAATAGAGCT----- 344
QY      61 AsnPheGlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGlu 80
DB      345 -----GGTGAAGACCAAACTAAT 362
QY      81 LysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThr 100
DB      363 AATGACGGAAATTTTCTGTACGTCCTAATTTGGCTACTCCATCTGGAAGACGGACG 422
QY      101 ValSerSerGluLysLeuValTyrAspValValIleuLysGlnAlaAlaLeuValLysArg 120
DB      423 ATGACATCGGAACACATGCTCATATGATGTGTTGAGCGACGACCTTGTGTAAGAG 482
QY      121 GlnLeuArgSerThrAspAspLeuGluValLysProAspIleValProGluValLeu 140
DB      483 CAACGAGATCTACCAATGATGATGAGAGGACCGGATATACCTATTCGGGGAATTTG 542
QY      141 GlyLeuLeuSerGluAlaTyrAspArgCysGlyGlyValCysAlaGluTyrAlaLysThr 160
DB      543 GCGTTGTGAGGAGCAATATCATGTATGTGTGTAAGATGTCAGAGATGCAAAAGACG 602
QY      161 PheTyrLeuGlyThrLysLeuMetThrProGluArgArgArgAlaIleThrAlaIleTyr 180
DB      603 TTTAATCTTNGAATCTATGCTAATGATCTCCGAGAGAAAGGCTATTCGGGCATATAT 662
QY      181 ValTTPCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrPro 200
DB      663 GTATGTCGAGAAAGACAGATGAACTTGTGTATGCCCCAAAGCATCATATTTACCCG 722
QY      201 GlnAlaLeuAspArgTTPGluThrArgLeuGluValIlePheSerGlyArgProPheAsp 220
DB      723 GCAGCTTAGATGATGGTGGGAAATAGGCTAGAGATGTTTCAATGGCGGCGCATTTGAC 782
QY      221 MetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPhe 240
DB      783 ATGCTCGATGCGCTTGTGCTTCATACATTTTCACTTCCAGTTGATATTCAGCCATT 842
QY      241 ArgAspMetIleGluGlyMetArgMetAspLeuTTPLysSerArgTyrLysThrPheAsp 260
DB      843 AGAGATATGATTTGAAGAGATGCGTATGACTTGAAAMAATGAGATACAAAACCTTCGAC 902
QY      261 GlnLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProVal 280
DB      903 GAACATATACCTTATGTTATATATGTTCTGCTAGCGTTGGGTATGATGATGCTTCCATT 962
QY      281 MetGlyIleAlaProGluSerLysValIleThrThrGluSerValTyrAsnAlaAlaLeuAla 300
DB      963 ATGGATATCGCCCTCAATCAAGGCAACACAGAGAGCGTATATATATGCTGCTTTGGCT 1022
QY      301 LeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArg 320
DB      1023 CTGGGGATCGCAAAATCAATTAATCAACATCTCAAGATGTTGGAAGAATGCCGAAGA 1082
QY      321 GlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIle 340
DB      1083 GGAAAGCTCTACTTGCCTCAAGATGAATTAGACACGGAGGCTATCCGATTAACATATATA 1142
QY      341 PheAlaGlyArgValThrAspLysThrArgAsnPheMetCysLysGlnIleGlnArgAla 360
DB      1143 TTTGCTGGAAGGGTACCGATTAATGAGAAATCTTTATGAGAAACCAATACATATGGGCA 1202
QY      361 ArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTTP 380
DB      1203 AGAAGGTTCTTTCATGAGCAGAGAGGAGCGGTGACAGAAATGAGCTCAGCTAGTATTC 1262

```

[illegible]

Db	543	GGCTTGTTGATGAGCATATGATAGGTGTGGTGAATGATAGTCGACGATATGCAAAAGACG	602
Qy	161	PheTyrLeuGlyThrIlysLeuMetThrProGluArgArgAlaIleTPrAlaIleTyr	180
Db	603	TTTAACTTAGGAACCTAAGCTAAATGACCTCCGAGAGAAAGGGGCTATCTGGCAATATAT	662
Qy	181	ValIleProValArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrPro	200
Db	663	GTAATGGCGAGAGAACAGATGAACTTTGTTATGGCCCAACCGCATATATATTAACCCCG	722
Qy	201	GlnAlaLeuAspArgTTPGJLurThrArgLeuGluAspIlePheSerGlyArgProPheAsp	220
Db	723	GCAGCCTTAGTATAGTGTGGAAAAATAGCCTGAAGATGTTTTCAATGGCGGCATTTGAC	782
Qy	221	MetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPhe	240
Db	783	ATGCTCCATGATGTGCTTTGTCCGATACAGTTTCAACTTTCAGTTGATATTCAGCCATTC	842
Qy	241	ArgAspMetIleGluGluMetArgMetAspLeuThrIlysSerArgTyrLvsThrPheAsp	260
Db	843	AGAAATATGATGTAAGGAATCGATAGACTTGAATAATCGAATACAAAACTTGAC	902
Qy	261	GluIleuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProVal	280
Db	903	GAACTATACCTTATATCTTATATATGTTGCTGGTACGGTTGGTTGATGAGTGTTCCAAAT	962
Qy	281	MetGlyIleAlaProGluSerLysValAlaThrThrGluSerValTyrAsnAlaAlaLeuAla	300
Db	963	ATGGGTATCCGCCCTGAAATCAAAAGCAACAACAGAGCCGATATATATATGCTGCTTGCGCT	1022
Qy	301	LeuGlyLeuAlaAsnGluLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArg	320
Db	1023	CTGGGGATCCGAAATCAATTAACCTACATCTCAGAAATGTGGAGAAAGATGCCAAGAGA	1082
Qy	321	GlyArgValTyrLeuProGluAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIle	340
Db	1083	GGAAGATCTTACTTGGCTCCTCAAGATGAAATTAGCAGCGCAGCTCATCTCGATGAAGATATA	1142
Qy	341	PheAlaGlyArgValThrAspLysTTPArgAsnPheMetLysGlnIleGlnArgAla	360
Db	1143	TTTCTCGAAGGGGTGACCGCATTAATGGAGAATCTTATGAAAGAAACAAATACATACAGGCA	1202
Qy	361	ArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTTP	380
Db	1203	AGAAAGTCTTTGATGAGCGAGAGAAAGCGGTGACAAATTAGCTACGCTAGTAGATTC	1262
Qy	381	ProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAsp	400
Db	1263	CTGTATGGGCATCTTTGCTCTGTACCGCAAAATACTAGATGAGATTGAAGCCAAATGAC	1322
Qy	401	TyrAsnAsnPheThrArgArgAlaTyrValSerLysProLys-LysLeuLeuThrLeuPr	420
Db	1323	TACAAACAACCTTCAACAAAGAGAGCATATGTAGCAAAATCAAAAGCAAGTTGATTCATTAC	1382
Qy	420	AlaIleAlaTyrAlaLysSerLeuValProProAsnArgThrSerSerProLeuAlaLys	439
Db	1383	TATTGCATATGCAAAATCTCTGTGCTCTCT-----ACAAACTGCTCTCTTCCAA	1434
RESULT 13			
LERYGTONS			
LOCUS	1355 bp	mRNA	linear
DEFINITION	L.esculentum (ry mutant) GTOWS mRNA for mutant phytoene synthase.		
ACCESSION	X67144		
VERSION	X67144.1 GI:19346		
KEYWORDS	GTOWS gene; mutant; phytoene synthase.		
SOURCE	Lycopersicon esculentum (tomato)		
ORGANISM	Lycopersicon esculentum		
REFERENCE	1 (bases 1 to 1355)		
AUTHORS	Fraty,R.G. and Grierson,D.		

TITLE Identification and genetic analysis of normal and mutant phytoene synthase genes of tomato by sequencing, complementation and co-suppression

JOURNAL Plant Mol. Biol. 22 (4), 589-602 (1993)

MEDLINE 93344508

PUBMED 8343597

REFERENCE 2 (bases 1 to 1355)

AUTHORS Fray, R.G.

TITLE Direct Submission

JOURNAL Submitted (01-JUL-1992) R.G. Fray, University of Nottingham, School of Agriculture, Sutton Bonington, Loughborough, Leicestershire LE12 5RD, UK

FEATURES

source Location/Qualifiers

1..1355

/organism="Lycopersicon esculentum"

/mol\_type="mRNA"

/strain="Alisa Craig"

/isolate="ry mutant"

/db\_xref="taxon:4081"

/chromosome="3"

/clone="p5ry"

/cissue\_type="pericarp"

/dev\_stage="ripe fruit"

1..1355

/gene="G5M5"

9..1307

/gene="G5M5"

/codon\_start=1

/product="mutant phytoene synthase"

/protein\_id="CAA47625.1"

/db\_xref="GI:19347"

/db\_xref="SPTREMBL:Q40166"

/translation="MSVALMTVSPDCVDSNGTSPMSVEBGRPFQSSPHRLVSENER  
INRGKQNNRKSPSVSAILATPSGETMTSEOMVYVTRQALVYRQLRSTNEL  
EKRPDIPFGNIGLISEAYDRGCEVCAEAKTPIKGTMLTERRALVATYWCRT  
DELVDGNASVITTPALDWEENLEEDVEGRPFMDLDGLSDVSNFPVDIQPFDM  
EEMRDLKRSRKNPDELTYCYVAVGTVGLMSVPMGLAPSKATTESVVAALAG  
IANQLTNLRDVGEDARGRVYLPDELDAGLSDEDI PAGVTDKMRIFMKQIHR  
KPFDEAKGVTELSASRPFVWASLVYRKILDEIDVANNFTKRAVVSNNMLKD  
PESNFKSGRSGNATTVLGLAPCEYIAIDRGPIGRTF"

variation

1171..1335

/gene="G5M5"

BASE COUNT 412 a 223 c 354 g 366 t

ORIGIN

Alignment Scores:

Pred. No.: 2.07e-134 Length: 1355

Score: 1668.00 Matches: 329

Percent Similarity: 86.27% Conservative: 29

Best Local Similarity: 79.28% Mismatches: 33

Query Match: 73.48% Indels: 24

DB: 8 Gaps: 3

US-09-847-081b-2 (1-440) x LERRYGTOM5 (1-1355)

QY 2 SerMetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThr 21

DB 6 TCATGTCCTGCTTCCCTTGTATGCGTTGTTCTCTCT--TGTGACCTCTCAAAATGGGACA 62

QY 22 GlyLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAla 41

DB 63 AGTTTCATGGAATCACTCCGGAGGGAACCGTTTTTTTGATTCAATCG-----110

QY 42 ArgAspArgAsnLeuMetTrpAsnGlyArgGlyLeuGlyGlyArgGlnArgTrpAsn 61

DB 111 AGGCATAGGAATTTGGTCCCAATGAGAGATCAATCAATAGAGT-----152

QY 62 PheGlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluLys 81

DB 153 -----GGTGGAAAGCAAACTAATAT 173

QY 82 GlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyLeuMetThrVal 101

DB 174 GGACGGAATTTTCTGACGGTCTGCTATTTTGGCTACTCATCTGAGAAAGGACATG 233

QY 102 SerSerGluLysValTyrAspValValLeuValGlnAlaLeuValLysArgGln 121

DB 234 ACATCGAACAGATGCTATGATGATGGTTTGAAGGAGGAGCGCTTGATGAAGGACA 293

QY 122 LeuArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeuGly 141

DB 294 CTGACATCTACCAATAGAGTTAGAAAGTGAAGCGGATATACCATTCGGGGGAATTTGGGC 353

QY 142 LeuLeuSerGluAlaTyrAspArgCysGlyGlyValGlyValAlaGlyTyrAlaLysThrPhe 161

DB 354 TTGTTGAGTGAAGCATATGATGATGCTGTGTGAAGTATGTCAGAGTATGCAAAAGACGTTT 413

QY 162 TyrLeuGlyThrIleLeuMetThrProGluArgGluAlaIleThrAlaIleTyrVal 181

DB 414 AACCTAGGAACATGCTAATGACTCCAGAGAAAGGCGCTATCTGGGCAATATATGTA 473

QY 182 TrpCysArgTyrThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGln 201

DB 474 TGTGCAAGAAACAGATGAACCTTGTGATGGCCAAAGCATCATATATATCCCGGCA 533

QY 202 AlaLeuAspArgTyrGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMet 221

DB 534 GCCTTAGATGAGTGGGAAATAGGCTAGAGATGTTTCAATGGCGGCATTGTGACATG 593

QY 222 LeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGluProPheArg 241

DB 594 CTCGATGGGCTTGTTCGATACAGTTTCTTAACITTCAGATTAATTCAGCATTCACA 653

QY 242 AspMetIleGluGlyMetArgMetAspLeuTyrLysSerArgTyrLysThrPheAspGlu 261

DB 654 GATATGATGGAAGAGATCGTATGACCTTGAGAAATGCAAGATACAAAACTCCAGACA 713

QY 262 LeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMet 281

DB 714 CATATCCTTATTTGTTATTAATCTTCTCGTACGTTGGTGTGATGAGTTCATATATG 773

QY 282 GlyIleAlaProGluSerIleValAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeu 301

DB 774 GGTATCGCCCTGAAATCAAGGCAACACAGAGACGATATTAATGCTGCTTGGCTCG 833

QY 302 GlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGly 321

DB 834 GGGATCGCAATCAATTAATCACTACACTACAGATGTTTGAGAAAGATCCGAAAGAGCA 893

QY 322 ArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePhe 341

DB 894 AGAGTCTACTTCCCTCAAGATGAATTAACACAGGAGGCTATCCGATGAAGATATATT 953

QY 342 AlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArgAlaArg 361

DB 954 GCTGGAAGGTCACCAATAAATGAGAAATCTTTATGAAAGAAACAAATCATATGGGCAAGA 1013

QY 362 LysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpPro 381

DB 1014 AAGTCTTGTATGAGGACAGAAAGGCGGTGACAGAAATGAGCTACGCTATGATTCCTT 1073

QY 382 ValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAspTyr 401

DB 1074 GTATGGGACATCTTGTGTCGTCGCAAAATACTAGATGAGATGGAAGCCATGACTAC 1133

QY 402 AsnAsnPheThrArgArgAlaTyrValSerLysProLysLysLeu 416

DB 1134 AACCACTTCACAAAGAGACATATGTGCAAAATATATGCTC 1178

RESULT 14

AF220218 1506 bp mRNA linear PLN 11-FEB-2000

DEFINITION Citrus unshiu phytoene synthase (Pey1) mRNA, complete cds.

ACCESSION AF220218

VERSION AF220218.1 GI:6959859

KEYWORDS



SOURCE  
ORGANISM Citrus unshiu  
Citrus unshiu  
Eudaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Sapindales; Rutaceae; Citrus.

REFERENCE  
TITLE Kim, I.-J., Ko, K.-C., Kim, C.-S. and Chung, W.-I.  
JOURNAL Isolation of a cDNA encoding phycoene synthase from Citrus  
AUTHORS unpublished  
REFERENCE 2 (bases 1 to 1506)  
TITLE Kim, I.-J., Ko, K.-C., Kim, C.-S. and Chung, W.-I.  
JOURNAL Direct Submission  
AUTHORS Submitted (30-Dec-1999) Biological Sciences, Korea Advanced  
JOURNAL Institute of Science and Technology, 373-1 Kusong-dong, Yuseong-gu,  
Taejon 305-701, South Korea

FEATURES  
source  
Location/Qualifiers  
1..1506  
/organism="Citrus unshiu"  
/mol\_type="mRNA"  
/culti\_var="Miyagawa (early)"  
/db\_xref="taxon:55188"  
/tissue\_type="fruit"  
/note="authority: Citrus unshiu (Swingle) Marc."  
1..1506  
/gene="psyl"  
112..1422  
/gene="psyl"  
/codon\_start=1  
/product="phycoene synthase"  
/protein\_id="AF33237.1"  
/db\_xref="GI:6959860"  
/translation="MSVTLLWVSPNSQSCFGVDSVREENLFYSRFLYQHOT  
TAVNRSKOFNNSKORNSYPLDTLRHSCSGIDLPISCMVATGAVMSSE  
MVYVVKQALVANKOPSGVRDLVDNDIALPGTSLISAVYRCGEVCAEYKTFY  
LGTLLMTERRAITAIYVWCRBDLYDGNASHITPTALDRRESRIEDFRPRPD  
MLDALSTVTKRPVDIOPFRDMLEGKMDLRKSKYKFDLYLYCYVQDELGLMSY  
PVMGLAPSOATTESVVAALAGLNDLTNLRVGEDARGRVYLPDELTAAGLS  
DDDIPEGVITIKRMFMNQIKRAMPEFDMANGVTLESEASRPVMSLILYRQILD  
EIEANDYVNFKRAVSRKAKKIALPIVAKSLRPSRIYSKA"

BASE COUNT 448 a 271 c 360 g 427 t

ORIGIN

Alignment Scores:  
Pred. Nc.: 9.9e-131 Length: 1506  
Score: 1626.00 Matches: 327  
Percent Similarity: 82.61% Conservative: 34  
Best Local Similarity: 74.83% Mismatches: 64  
Query Match: 71.63% Indels: 12  
Gaps: 4

US-09-847-081b-2 (1-440) x AF220218 (1-1506)

QY 2 SerMetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThr 21  
Db 109 AGCATGTCGTGTACATTCTGTGGGTGTATCACTCACTCAATTCGATTCCTTC 168  
QY 22 GluLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40  
Db 169 GGGTTCGTCGATTCAGTTCGAGAGAAACAGCGCTGTTTATTCTATCAAGATTTCCTTAC 228  
QY 41 --AAlaArgAspArgAsnLeuMetTrpAsnGlyArgIleGlyLys-----GlyGly 56  
Db 229 CAACATCAAAACCGGACCTGCTGTATTATTCTAAGCACTTAATATATAGTAT 288  
QY 57 ArgGlnArgTrpAsnPheGlySerLeuIleAlaAspProArgTrpSerCysLeuGlyGly 76  
Db 289 AAGCAGAACCGAATTCCTTATCTTATAGATACAGATTGAGGACATCTTGC----- 339  
QY 77 SerArgThrGluGlySerThrPheSerValGlnSerSerLeuValAlaSerProAla 96  
Db 340 -----TATCTGGAATCACTTCCTGTAATAATATCATGTATGCTTACGACCTCT 390  
QY 97 GlyGluMetThrValSerSerGluGlyLysValTyrAspValValLeuLysGlnAlaAla 116

Db 391 GGAGAAATGGCCATGCTTCAGAAAGAAATGTTTACATATGTTGCTCAAGCAGGCAGCC 450  
QY 117 LeuValLysArgGluLeuAsnSer---ThrAspAsnLeuGluValLysProAspIleVal 135  
Db 451 TTGGTATATAGAACCAAGTGGGGTTTACCTGATATCTTATGTAAGCAACGATATTTGCT 510  
QY 136 ValProGlyAsnLeuGlyLeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAla 155  
Db 511 TTACCCGGAACCTTAATCTGCTCAGTGAAGCTTATATATGTTGTGAGAAATTTGGGCC 570  
QY 156 GluTyrAlaLysThrPheTyrLeuGlyThrLysLeuMetThrProGluArgArgAla 175  
Db 571 GAGTATCTAAGACATTTTACTTGGGAACTTTGCTGATGACCTCTGAAAGCGGAAGGCGCT 630  
QY 176 IleTrpAlaIleTyrValITrPAspArgTrpAspGluLeuValAspGlyProAsnAla 195  
Db 631 ATATGGGCTATATATGTGTGTGTAGAGAGACATATAGCTCGTTGATGGGCTTAATGCT 690  
QY 196 SerHisIleThrProGlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSer 215  
Db 691 TCACACATTACTCCAACAGCTTTAGACAGGTGGAGTCCAGGTTGGAAGACCTTTCCGG 750  
QY 216 GlyArgProPheAspMetLeuAspAlaAlaLeuSerAspThrValSerArgPheProVal 235  
Db 751 GGTGTCATTTGATGATGCTTGAATGCTGATTCATATCAGATACAGTAAATTTCTGTC 810  
QY 236 AspIleGlnProPheArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArg 255  
Db 811 GACATTCAGCCATTCAGAGATATGATATGAAGATGAGATGAGACCTTAGAATCTAAGA 870  
QY 256 TyrIleThrPheAspGluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeu 275  
Db 871 TACAAATCTTATGATGATTAATCTGATATGTTATATATGTTGTGGACCGTAGGGCTA 930  
QY 276 MetSerValProValMetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyr 295  
Db 931 ATGAGTTCACAGTTATGGGCTATGAGCATGACCTGACCTCAACAGCAACAGAGAGGCTTAC 990  
QY 296 AsnAlaIleLeuAlaLeuGlyLeuAlaAsnGluLeuThrAsnIleLeuArgAspValGly 315  
Db 991 AATGACACATTTGGCATAGGATTTGCTAATCACTCACTCACTCACTCACTCACTCACTCA 1050  
QY 316 GluAspAlaArgAspGlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeu 335  
Db 1051 GAGGATCCCGGAAGAGAGAGGTTTATCTACCAAGATAGTTGGCAGCAGGCGGCTT 1110  
QY 336 SerAspGluAspIlePheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysLys 355  
Db 1111 TCAGATGATGACATATTTGCTGAGAGAGTGACCATTAATGAGAAACCTTCATGAAAGAAC 1170  
QY 356 GlnIleGlnArgAlaArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAsp 375  
Db 1171 CAATTTAAGAGGAGAGAGATTTTGTGATATGCTGTAAGAACGATGACCGAGCTTACT 1230  
QY 376 SerAlaSerArgTrpProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGlu 395  
Db 1231 GAACATGTCATATGCGCGGTATGCGCTCATCTGTTGTATACGGCAATATCTGATGAG 1290  
QY 396 IleGlnAlaAsnAspTyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysLys 415  
Db 1291 ATTGAGGCCAATGATTCACCAACTTCAAAAGAGAGACCTTATGATGAATAAGCCAAAG 1350  
QY 416 LeuLeuThrLeuProIleAlaTyrAlaLysSerLeuValProProAsnArg 432  
Db 1351 ATAGCTCACATTCACATTTGATGCAAAATTCCTTTAGCCCGCTGAAGA 1401

RESULT 15  
AB037975 1689 bp mRNA linear PLN 25-NOV-2000  
DEFINITION Citrus unshiu mRNA for phycoene synthase, complete cds.  
ACCESSION AB037975  
VERSION AB037975.1 GI:11344506



KEYWORDS phytoene synthase.  
SOURCE Citrus unshiu  
ORGANISM Citrus unshiu  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Sapindales; Rutaceae; Citrus.  
REFERENCE 1 (sites)  
AUTHORS Ikoma,Y., Komatsu,A., Kita,M., Ogawa,K., Omura,M., Yano,M. and  
Moriguchi,T.  
TITLE Expression of a phytoene synthase gene and characteristic carotenoid  
accumulation during citrus fruit development  
JOURNAL Physiol. Plantarum 111, 232-238 (2001)  
REFERENCE 2 (bases 1 to 1689)  
AUTHORS Moriguchi,T.  
TITLE Direct Submission  
JOURNAL Submitted (04-FEB-2000) Takaya Moriguchi, National Institute of  
Fruit Tree Science, Department of Research Planning and  
Coordination, 2-1 Fujimoto, Tsukuba, Ibaraki 305-8605, Japan  
(E-mail: takaya@fruit.affrc.go.jp, Tel: 81-298-38-6416,  
Fax: 81-298-38-6437)  
FEATURES  
source location/Qualifiers  
1..1689  
/organism="Citrus unshiu"  
/mol\_type="mRNA"  
/cultivar="Satsuma mandarin"  
/db\_xref="taxon:55188"  
/cissue\_type="pulp"  
/dev\_stage="mature stage"  
179..1489  
/codon\_start=1  
/product="phytoene synthase"  
/protein\_id="BAB18514.1"  
/db\_xref="GI:11344507"  
translation="MSVTLLWVSPNSOLNCGFVDSYRENRLEFYSRPIYCHOTR  
TAVFSPKQFNKSNKQRNSPYLDLDRHPSGSDIDPEICMTASTRGRVMSSE  
MYNVYVKQALVNVKQFSGVTRDLVNPVRIPLGTLISEAYDCEGVCAVATFY  
LGLMTSERRRAIVAIYWCRTDELVDGPNASHITPVALDRMWSRLDELFRGPF  
MLDALSDVTVEPVDIQFMDIEGMRDLKSKYKNEDELYLYCYVAGVGLMSV  
PVMGIAPDSQATFESVYNALALGIANOLTLRDVGEDARGRVYLPDDELQAGLS  
DDIDPAGEVTITKRMFMKNOIKRAMPFPMANGVTLEISEARWPMASLLYRQILD  
ELEANDYNNFTGRAVSAKAKTAAPIAIVAKSLRPSRLYTSKA"  
BASE COUNT 512 a 302 c 368 g 487 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,15e-130 Length: 1689  
Score: 1626.00 Matches: 327  
Percent Similarity: 82.61% Conservative: 34  
Best Local Similarity: 74.83% Mismatches: 64  
Query Match: 71.63% Indels: 12  
DB: 8 Gaps: 4  
US-09-847-081b-2 (1-440) x AB037975 (1-1689)  
QY 2 SerMetSerValAlaLeuLeuTrpValAlaSerProThrSerGluValSerAnGlyThr 21  
DB 176 AGCAGTCTGTTACATTGCTGTGGTGTATCATCACTCAATTCACATTCGCTTC 235  
QY 22 GlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40  
DB 236 GGGTTCGTCGATTCAGTTCGAGAGAAAACAGCGTGTATTATTCATCAAGATTTCTTTAC 295  
QY 41 ---AlaArgAspArgAsnLeuMetTrpAsnGlyArgGlyLeuValys-----GlyGly 56  
DB 236 CAACATCAAAACCCGACGCTGTGTTAATTTCTAGACCTTAACAGATTATATATGTAAT 355  
QY 57 ArgGlnArgTrpAsnPheGlySerLeuLeuAlaAspProArgTrpSerCysLeuGlyGly 76  
DB 356 AAGCAGAGACGGAATTCATTATCTTATAGATACAGATTAGAGCATCTTC----- 406  
QY 77 SerArgThrGluGlyGlySerThrPheSerValGlnSerSerLeuValAlaSerProAla 96  
DB 407 -----TCATCTGGAATCGACTTGCCTGAATAATATCATGTATGCTGTGCTAGCACTGCT 457

QY 97 GlyGluMetThrValSerSerGluTrpSerValIleAspValIleLeuLysGlnAla 116  
DB 458 GGAGAGAGGCGCATCTCTTCAGAAAGAAATGCTTACATGTTGCTGCTCAGACGACGCC 517  
QY 117 LeuValIleArgGlnLeuAspSer--ThrAspAspLeuGluValIleProAspIleVal 135  
DB 518 TTGGTTAATATGCAACCAAGTGGGCTTACTCGTGAATCTTGAATGACCAAGATATTGCT 577  
QY 136 ValProGlyAsnLeuGlyLeuLeuSerGluAlaIleAspArgCysGlyGluValCysAla 155  
DB 578 TTACCCGGAACCTTAAGTCTGCTCAGTAAGCTTAATGATCGTTGGAGAAATTTGGCCG 637  
QY 156 GlyTrpAlaIleArgThrPheTrpLeuGlyThrIleLeuMetThrProGluArgArgAla 175  
DB 638 GAGTATGCTAAGACATTTTACTTGGGAACTTTCGTCATGACCTCTGAAAGGGAAGCGCT 697  
QY 176 IleTrpAlaIleIleTrpValIleTrpCysArgArgThrAspGluLeuValAspGlyProAsnAla 195  
DB 698 ATATGGCTATATATATGTGTGTGTGTGAGAGACAGATGAGCTGTTGATGGCCTATGCT 757  
QY 196 SerHisIleThrProGlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSer 215  
DB 758 TCACACATTAACCTCAACAGCTTTAGACAGGTGGAGTCCAGGTGGAAGACCTTTTCCGG 817  
QY 216 GlyArgProPheAspMetLeuAspAlaAlaLeuSerSerAspThrValSerArgPheProVal 235  
DB 818 GGTCTCCATTTGATATGATCTGATGCTGCTGATATATAGATACAGTAAACCAATTTCTGTC 877  
QY 236 AspIleGlnProPheAspMetLeuGluGlyMetArgMetAspLeuTrpLysSerArg 255  
DB 878 GACATTCAGCATTGACAGATATGATGATGAGAAATGAGATGAGACCTTAGAAGTCAAGA 937  
QY 256 TyrIleSerThrPheAspGluLeuIleTrpLeuTrpCysTrpTrpValAlaGlyThrValGlyLeu 275  
DB 938 TACAAAAACCTTGAAGAAATTATATCTGTATTATATTATGTTGCTGGAGCCGTAGGCTCA 997  
QY 276 MetSerValProValMetGlyIleAlaProGluSerIleAlaThrGluSerValIleTrp 295  
DB 998 ATGAGTGTTCAGATTATGAGCATGACCTGATCAGGCAACAGAGAGCTGTC 1057  
QY 296 AsnAlaAlaLeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuAspArgValGly 315  
DB 1058 AATGCAGATTGGCAGCTAGGAGTGTCTATATGCTATACGCTCACTAATCATCTCAGAGATGTTGA 1117  
QY 316 GluAspAlaArgArgGlyArgValIleTrpProGlnAspGluLeuAlaGlnAlaGlyLeu 335  
DB 1118 GAGGATGCCCAAGAGGAGGAGGATTATCTTACACAAAGATGAGTTGGCACAGGACGGCTT 1177  
QY 336 SerAspGluAspIlePheAlaGlyValArgValIleAspIleSerThrAspAsnPheMetLysLys 355  
DB 1178 TCAGATGATGACATATTTGCTGAGAGAGTACCATTAATATGAGAACTTCATGAGAAC 1237  
QY 356 GlnIleGlnArgAlaArgLysPhePheAspGluSerGluLysGlyValIleThrGluLeuAsp 375  
DB 1238 CAATTAAGAGAGGCAAGATGTTCTTATATGCTGAGGAACGGTGTACCGAGCTGAGT 1297  
QY 376 SerAlaSerArgTrpProValLeuThrAlaLeuLeuLeuTrpArgLysIleLeuAspGlu 395  
DB 1298 GAAGCTATGATGAGCGCGATGAGGCTTCATTCGTTGTCGACCGCAATATCTGATAG 1357  
QY 396 IleGluAlaAsnAspTrpAsnAsnPheThrArgArgAlaIleValSerLysPheProLysLys 415  
DB 1358 ATTGAGGCAATGATTAACAACAATCTTCACAAAGAGCTTATGTAGTAAGAACCAAGANG 1417  
QY 416 LeuLeuThrLeuProIleAlaIleValIleValSerLeuValProProAsnArg 432  
DB 1418 ATAGCTGACATCAATGACATGCAAAATCCCTTACGCCGCTCAAGA 1468

Search completed: January 16, 2004, 08:00:26  
Job time : 5065 secs





```

FT      /tag= a
FT      /product= "phytoene synthase"
XX      DE10022362-A1.
XX      15-NOV-2001.
XX      08-MAY-2000; 2000DE-1022362.
XX      08-MAY-2000; 2000DE-1022362.
XX      (PARB ) BAYER AG.
XX      Buech M, Hain R;
XX      MPI; 2002-027336/04.
XX      P-P8DB; AAM51841.
XX      New nucleic acid encoding tobacco zeta-carotene desaturase, useful for
XX      screening compounds with herbicidal activity -
XX      Claim 14; Page 12-17; 44pp; German.
XX      The present invention provides the protein and coding sequences of
XX      phytoene synthases and zeta-carotene desaturase from Nicotiana tabacum.
XX      The sequences can be used to identify compounds capable of altering the
XX      expression of these genes, which are therefore useful as plant growth
XX      regulators and herbicides. They can also be used to produce transgenic
XX      plants. The present sequence is the coding sequence of a tobacco
XX      phytoene synthase.
XX      Sequence 1728 BP; 513 A; 290 C; 433 G; 492 T; 0 other;
SQ      Alignment Scores:
Pred. No.:      5,6e-234      Length:      1728
Score:          2270.00      Matches:      440
Percent Similarity: 100.00%      Conservat: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:            24      Gaps:      0
US-09-847-081B-2 (1-440) x AAI66366 (1-1728)
Oy      1 MetSerMetSerValAlaLeuLeuThrValValSerProThrSerGluValSerAnsgly 20
Db      244 ATGACACATCTCTGTTGCTTTGTTGGGTGTTTCTCCACCTCCGAGGCTCCGAAATGG 303
Oy      21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40
Db      304 ACAGGATGTTGGATTTCAGTCCGAGAGAGAAACCGGTCCTTGTATCATCCAGGTTCTTA 363
Oy      41 AlaArgAspArgAsnLeuMetTrpAsnGlyArgGlyLeuGlyValArgGlnArgTrp 60
Db      364 GCTCAGATGAGAAATTTGATGTGGAATGGAGAAATCAAGAAAGGTGGAGCAAAAGGTGG 423
Oy      61 AsnPheGlySerLeuIleAlaAspProArgTrpSerCysLeuGlyGlySerArgThrGlu 80
Db      424 AATTTTGGCTCTTAATTCCTGATCCAAATATTATGCTTGGGATGATCAAGAACTGAA 483
Oy      81 LysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThr 100
Db      484 AAGGAGACACTTCTCTGTACAGTCCAGTTGGTGGTAGCCAGCTCGAGAAATGACT 543
Oy      101 ValSerSerGluValGlyValTyrAspValValLeuGlyGlnAlaLeuValIleArg 120
Db      544 GTGTATCAGAGAAAGAGTGTATGTATTAAGCGGCGCTTTAGTGAAGAG 603
Oy      121 GlnLeuArgSerThrAspAspLeuGluValLysProAspIleValProGlyAsnLeu 140
Db      604 CAGCTGAGATCTACCGATGATTTAAGATGAAGCCGGAATATTGTTCCAGGGAAATTGG 663
Oy      141 GlyLeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLeuThr 160

```

```

Db      664 GCCTTGTGAGTGAACATATGATCGTTGGCGAAGTATGTGACATATGCAAGACA 723
Oy      161 PheTyrLeuGlyThrIleLeuMetThrProGluArgArgAlaIleTyrPalaIleTyr 180
Db      724 TTTTACTTAGAACCAAGCTTATATGACCCAGAGAGAAAGAGCTTATCTGGCAATATAT 783
Oy      181 ValTrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrPro 200
Db      784 GTGTGTGACAGAGAACGATGATGACTGTTGATGGCCCTTAATGATCCACATATACCTCG 843
Oy      201 GlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAsp 220
Db      844 CAAGCTTTAGATGAGGAGACCAAGCTGGAAGATATTTTCAAGTGGCGCCCATTTGAT 903
Oy      221 MetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPhe 240
Db      904 ATGCTTATGCTGCTTTATTCGATCTGTCTCCGATTTTCTGTGATATTCAGCCATTTC 963
Oy      241 ArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrIleThrPheAsp 260
Db      964 AGAGATATGATGAGAGAAATGCGTATGACCTTGAGAAATCCAGATACAAAACTTTCGAT 1023
Oy      261 GluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProVal 280
Db      1024 GAGCTATATCTCTATTTGTTACTATGTGCTGTGCTAGTCTAGATTTGATGAGAGTTCCAGTT 1083
Oy      281 MetGlyIleAlaProGluSerLysAlaThrThrgluSerValTyrAsnAlaAlaLeuAla 300
Db      1084 ATGGGATTTGCACTTGATCAAGATCAAGCAACAGAGAGTGTATTTATGCTGCTTGGCT 1143
Oy      301 LeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArg 320
Db      1144 TTAGGCTTGGAATCAACTATACCAATATATCTGAGATGATGAGAGAAATGCCAGAGA 1203
Oy      321 GlyArgValTyrLeuProGluAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIle 340
Db      1204 GGAAGATATATACCTGCTCAAGATGATTTGACACAGGCAAGGCTCTCCGAGAGACATA 1263
Oy      341 PheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArgAla 360
Db      1264 TTTCCTGAAAGAGTGAATGATTAAGTGAAGAACTTTATGAAAGAAATTCAGAGGGCG 1323
Oy      361 ArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrp 380
Db      1324 AGGAAATTTCTTATATGATGACAGAAAGGTGTACACAACTGCACTGCTAGTAAATGG 1383
Oy      381 ProValLeuThrAlaLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAsp 400
Db      1384 CCTGTGTTAAACAGGCTGCTGTGTATCGCAAGATATTTGACGAGATTGAAGCCAAACGAC 1443
Oy      401 TyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuPro 420
Db      1444 TACAACTTTCACAAAGAGGCGCTTATGTATGCAAGCAAGAAAGCTTTCACCTTCCC 1503
Oy      421 IleAlaTyrAlaLysSerLeuValProProAsnArgThrSerSerProLeuAlaLysThr 440
Db      1504 ATTGCTATGCAAAATCTCTTGTGCCCCCTAATATGAACCTTCTCCACTAGCAAAACA 1563

```

RESULT 2  
 ABE13568 standard; DNA; 1573 BP.  
 ID ABE13568;  
 XX ABE13568;  
 AC  
 XX  
 DT 12-JUN-2003 (first entry)  
 XX  
 DE Tobacco phytoene synthase DNA.  
 XX  
 KW Zeaxanthin epoxidase; zeaxanthin; carotenoid; inhibitor;  
 KW transgenic plant cell; alpha-tocopherol; food supplementation;  
 KW tobacco; phytoene synthase; Fys; ds.  
 XX  
 OS Nicotiana tabacum.

XX WO2002103021-A2.  
 XX 27-DEC-2002.  
 XX 19-JUN-2002; 2002WO-BP06810.  
 XX 19-JUN-2001; 2001EP-0114661.  
 XX (VKAM/) VON KAMEKE K.  
 XX Sandmann G, Roemer S, Luebeck J, Adomat C, Kauder F;  
 XX WPI; 2003-167525/16.  
 DR Increasing zeaxanthin and total carotenoid contents of plants, useful  
 PT for food supplementation, by inhibiting activity of zeaxanthin  
 PT epoxidase  
 PT  
 PS Examples; Page 34-35; 52pp; German.  
 XX  
 CC This invention describes a novel method of increasing the zeaxanthin  
 CC and/or total carotenoid contents in transgenic plant cells, plants or  
 CC other host cells or organisms (e.g. algae) by inhibiting the endogenous  
 CC zeaxanthin epoxidase (ZEP) activity. The invention also discloses  
 CC transgenic plant cells or plants, their harvested products, replicative  
 CC material (protoplasts, calli, seeds, tubers and cuttings) and  
 CC descendants, produced by the novel method. The method is used to increase  
 CC production of zeaxanthin and total carotenoids (including the therapeutic  
 CC antioxidant alpha-tocopherol) for food supplementation. This sequence  
 CC represents a polynucleotide corresponding to the tobacco (Nicotiana  
 CC tabacum) phytylene synthase (Pys) gene which is used to describe the  
 CC method of the invention.  
 XX  
 SQ Sequence 1573 BP; 475 A; 270 C; 390 G; 432 T; 6 other;

## Alignment Scores:

Pred. No.: 26-220 Length: 1573  
 Score: 2143.50 Matches: 425  
 Percent Similarity: 96.83% Conservative: 3  
 Best Local Similarity: 96.15% Mismatches: 12  
 Query Match: 94.43% Indels: 3  
 DB: 25 Gaps: 1

US-09-847-081B-2 (1-440) x ABX13568 (1-1573)

QY 1 MetSerMetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerangly 20  
 DB 10 ATGACCATGTCCTGTTGCTTGTGGATGTTTCCCAATTCGAGGTCCTCAATGGG 69  
 QY 21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40  
 DB 70 ACGAGATCTTGTGATTCAGTCGAGATGGAACCGGCTTTGTATCATCCAGGTTCTTA 129  
 QY 41 AlaArgAspArgAsnLeuMetTrpAsnGlyArgGlyLeuGlyGly---ArgGlnArg 59  
 DB 130 GCTCGACGTAGAGATTGATGTGGAATGGAGATCAAGAAAGCCGGAGAGACAAAG 189  
 QY 60 TrpAsnPheGlySerLeuLeuAlaAspProArgTrpSerCysLeuGlyGlySerArgThr 79  
 DB 190 TGGAAATTTGGCTCTTAATTCAGATTCAGATATGATGCTTGGNNNGATCAAGAACT 249  
 QY 80 GluLysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMet 99  
 DB 250 GANNNTGAGAACCTTCTCTGACAGTCAGTTGGTGGCTAGCCACCTGAGAAAG 309  
 QY 100 ThrValSerSerGlyLysLysValTyrAspValValLeuLeuGlnAlaLeuValLys 119  
 DB 310 ACTGTCTCATCAGAAAGAAAGGTATGTATGTGATTTAAAGCAGCAGCTTTAGGAAG 369  
 QY 120 ArgGlnLeuArgSerThrAspAspLeuGluValLysProAspLleValValProGlyAsn 139  
 DB 370 AGCGAGCTGAGATCTTACCATGATTTAGAAAGCCGAGATATTGTTGTCACGAGAA 429

QY 140 LeuGlyLeuLeuSerGluAlaTyrAspArgCysGlyGlyValCysAlaGluTyrAlaLys 159  
 DB 430 TTGGGCTTGTTGAGTGAAGCATATGATGCTTGTGGCGAAGATGTGCGAGATGCAAG 489  
 QY 160 ThrPheTyrLeuGlyThrLysLeuMetThrProGluValArgArgAlaIleTrpAlaIle 179  
 DB 490 ACATTCTTACCTTAGGAACCAAGCTATATGACCCAGAGAAAGAAAGACTATCTGGCAATA 549  
 QY 180 TyrValTyrCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThr 199  
 DB 550 TATGTGTGTGAGAGAACCGATGAGCTGTGTGATGGCCCTTAATGCAATCCACATAACT 609  
 QY 200 ProGlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPhe 219  
 DB 610 CCGCAAGCTTTAGATAGTGGGAGACCGAGCTCGAAGATTA-MACAGTGGCGGCATT 668  
 QY 220 AspMetLeuAspAlaIleLeuSerAspThrValSerArgPheProValAspIleGlnPro 239  
 DB 669 GATATGCTTGATGCTGCTTATCCGATCTGCTCCAGATTTCTGTGATATTCAGCCA 728  
 QY 240 PheArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrIleThrPhe 259  
 DB 729 TTCAGAGATATGATTGAAGAGATGGATGACCTTGTGAAATCCAGATACAAAACCTTTC 788  
 QY 260 AspGluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValPro 279  
 DB 789 GATGAGCTATATCTTATATGTTTACTATGTTGCTGATCTGATGATTAATGAGTTCGA 848  
 QY 280 ValMetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaLeu 299  
 DB 849 GTTATGGTATATGCACTGATCAATCAAGCAACACAGAGAGTATATATGCTCTTGG 908  
 QY 300 AlaLeuGlyLeuAlaAsnGluLeuThrAsnIleLeuArgAspValGlyLysAspAlaArg 319  
 DB 909 ACTTTAGGCTTGCAATCAACTAACATATATCTCAGAGATGAGAGAGATGCCAGA 968  
 QY 320 ArgGlyArgValTyrLeuProGlnAspGluLeuAlaGlyLeuSerAspGluAsp 339  
 DB 969 AGAGAAAGATATCTTCTCCATGATATATAGACAGCGAGGCTCTCCGACCAAGAC 1028  
 QY 340 IlePheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysGlnIleGlnArg 359  
 DB 1029 ATATTGCTGGAAGAGTACATGATAGTGAAGAACTTTATGAAGAAACAATTGAGAG 1088  
 QY 360 AlaArgLysPhePheAspGluSerGlyLysValThrGluLeuAspSerAlaSerArg 379  
 DB 1089 GCGAGAAATCTTGTGATGAGTCAGAGAAAGGTGCACAGAACTGACTGCTGTGTAGA 1148  
 QY 380 TrpProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsn 399  
 DB 1149 TGCGCTGTGTAAACGCCCTGCTGTGTATGCAAGATATTTGAGAGATTTGAACCCAAC 1208  
 QY 400 AspTyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysLeuLeuThrLeu 419  
 DB 1209 GACTATCAACAACCTTCAAGAGAGGCTTATGTTGCAAGCCAAAGAACTTCTCACCTTG 1268  
 QY 420 ProIleAla-TyrAlaLysSerLeuValProProAsnArgThrSerSerProLeuAlaLys 439  
 DB 1269 CCCATTTGCTTATGCAAAATCTTTGTGCCCCCTTAATGAACTTCTCTCCACTGCAAA 1328  
 QY 439 sThr 440  
 DB 1329 GACA 1332  
 DB  
 RESULT 3  
 AA03880  
 ID AA03880 standard; cDNA; 1826 BP.  
 XX AA03880;  
 AC  
 XX  
 DT 29-APR-1998 (first entry)  
 XX

DE		Phytoene synthase coding sequence from N. tabacum.
XX		
KM		Phytoene synthase; transgenic plant; enhanced carotenoid synthesis;
KM		ultra violet absorber; food colour; ss.
XX		
OS	Nicotiana tabacum.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	367..1599
FT		/tag= a
XX		
PN	US5705624-A.	
XX		
PD	06-JAN-1998.	
XX		
PF	27-DEC-1995;	95US-0579667.
XX		
PR	27-DEC-1995;	95US-0579667.
XX		
PA	(IDELL/) DELLA-CIOPPA G R.	
PA	(FITZ/) FITZMAURICE W P.	
PA	(GRILL/) GRILL L K.	
PA	(HELL/) HELLMANN G M.	
PA	(KUMAI/) KUMAGAI M H.	
XX		
PI	Della-cioppa GR, Fitzmaurice WP, Grilli LK, Hellmann GM,	
P1	Kumagai MH;	
DR	WPI, 1998-086196/08.	
DR	P-PSDB; AAW41059.	
XX		
PT	DNA encoding tobacco phytoene synthase polypeptides - useful for	
PT	producing recombinant polypeptides or transgenic plants	
XX		
PS	Claim 1; Column 27-30; 25pp; English.	
CC		
CC	This sequence encodes the phytoene synthetase from Nicotiana tabacum.	
CC	The phytoene synthetase coding sequence represents a cDNA of the	
CC	invention. The isolated nucleic acid molecules are used for producing	
CC	recombinant polypeptides or transgenic plants with enhanced ability to	
CC	synthesize carotenoids. Phytoene has been used as a ultra violet absorber	
CC	and other carotenoids have been used as food colours, animal feeds and in	
CC	the pharmaceutical and cosmetics industries.	
XX		
SQ	Sequence 1826 BP; 589 A; 282 C; 433 G; 522 T; 0 other;	
	Alignment Scores:	
	Pred. No.: 3, 61e-186 Length: 1826	
	Score: 1826.00 Matches: 362	
	Percent Similarity: 88.61% Conservative: 27	
	Best Local Similarity: 82.46% Mismatches: 21	
	Query Match: 80.44% Indels: 29	
	DB: 19 Gaps: 4	
	US-09-847-081B-2 (1-440) x AAV03880 (1-1826)	
OY	3 MetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThrGly 22	
Db	367 ANGTCGTGGCTTGTATTGGATTTGTTCACT---TGTCAGAGTTCACAAATGGACAGGA 423	
OY	23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42	
Db	424 TTCTTGAGTTCAGTAGGAGGGAACCGGGTTTTGCATCCGCG-----AGG 471	
OY	43 AsparGanLeuMetTrpAsnGlyArgGlyIleValysGlyGlyValArgGlnArgTrpAsnPhe 62	
Db	472 CATGAGATTATGTGTGCATTCAGAACATCAAAGAAGGTGTGAACAAGAGTGAATTT 531	
OY	63 GlySerLeuIleAlaAspProArgTrpSerCysLeuGlyGlySerArgThrGluysGly 82	
Db	532 GGT----- 534	
OY	83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyLutMet---ThrVal 101	

Db	535	-----TCTGTACGCTCTGCCATGGTGGCTTACCAACGGGAGAAATGGCCACAAATG	585
Qy	102	SeSeSgcLulysylValTyrAspValValLeuylsglnAlaAleuVallyAspArgln	121
Db	586	ACATCAGAAACAGAGAGTTATGATGTGGTATTTGAAACACAGACGTTTAGTCAAAAGCAG	645
Qy	122	LeuArgSerThrAspAspLeuGlnValLysProAspIleValProGlyIaenLeuGly	141
Db	646	CTGAGAGTCTACGATGATATTAGAACTGAAGCCGGAGATCCTCTCCCGGGAAATTGAGC	705
Qy	142	LeuLeuSergLulAlatyAspArgCysGlyGlnValCysAlaGluTyrAlalyThrPhe	161
Db	706	TTCTTAGTGAAGACAAATGATAGGTGTAGTAAAGTATGCCAGAGTATGCCAAAGACATTT	765
Qy	162	TyrLeuGlyThrLysLeuMetThrProGluArgArgArgAlaIleTPrAlaIleTyrVal	181
Db	766	TACTTAGGAACATAGCTATGACTCCAGAGAGAAAGAGGCTTATTGGCAATATATATGTA	825
Qy	182	TrpCysArgArgThrAspGlnLeuValAspGlyProAlaIAspThrIleThrProGln	201
Db	826	TGGTGGCAGAGAAACAGATGAACCTGTATTATAGCCCGAATGATCACTATTATCCCAAA	885
Qy	202	AlAleuAspArgTrpGluThrArgLeuGlnLysIlePheSergLysArgProPheAspMet	221
Db	886	GCTTATGATAGGTGGGAGAACCGGCTGGAAGATGTTTCACTGGGCGGCCATTTGATATG	945
Qy	222	LeuAspAlaAleuSerAspThrValSerArgPheProValAspIleGlnProPheArg	241
Db	946	CTCGATGCTGCTTTGTCGATCTGTTCCAGGTTCCAGTTGATATTACGCGCTTACA	1000
Qy	242	AspMetIleGluGlyMetCArgMetAspLeuTrpLysSerArgTyrLysThrPheAspGlu	261
Db	1006	GAAATATATGAAGAAATGGTATGGACTTGAGGAAGTCAAAATACAAATCAAACTTTGATGAG	1066
Qy	262	LeuTyrLeuTyrCysEtyTyrValAlaGlyThrValGlyLeuMetSerValProValMet	281
Db	1066	CTTATACCTATATGTTTATTACGTTCTGCTGACGGTTGGGTTGATGATGTTCCAAATTATG	1122
Qy	282	GlyIleAlaProGlnSerLysAlaThrThrGlnSerValTyrAsnAlaAleuAlaLeu	301
Db	1126	GGTATTGCACTGATTCMAAGGCCAACACAGAGAGTGAATATATATGACGTTTGGCTTTA	1188
Qy	302	GlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGly	321
Db	1186	GGGATGCAAACTCAACTAACAACACTACTCAGAGATGTCGAGAAAGATGCCAGAAAGGA	1244
Qy	322	ArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePhe	341
Db	1246	AGAGTCTACTTACCTCAGATGATGTTAGACAGGAGGATCTCTCCGCAATGACATTTT	1300
Qy	342	AlaGlyArgValThrAspLysTrpArgPheMetLysLysGlnIleGlnArgAlaArg	361
Db	1306	GCTGGAAGAGTGATGAATAATGGAAGAGCTTTATGAAGAGCAAACTCAGAGGGCAAGA	1366
Qy	362	LysPhePheAspGlnSerGluGlyAlaThrGluLeuAspSerAlaSerArgTPrPro	381
Db	1366	AAATTTCTTGACAGGACAGACAGAGAGATGACACACACTGACCTCAGCTAGTAGTGCCCT	1422
Qy	382	ValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGlnAlaAsnArgTyr	401
Db	1426	GTAATGGCAATCTTGTGTGTGTAACGCCAGCAAGTACTCGACGATTTGAAGCCAAATGACTAC	1488
Qy	402	AsnAsnPheThrArgArgAlaTyrAlaSerLysProLysLysLeuThrLeuProIle	421
Db	1486	AAACAATCTCAACAGAGAGACTTATGTGACCAACCAAGAAAGCTATTTCTTACATAT	1544
Qy	422	AlaTyrAlaLysSerLeuValProProAsnArgThr-SerSerProLeuAlaLys	439
Db	1546	GCTTATGCAAAATCTCTGTGGCCCTCAAGAACTTGTGACCTCTAGCTAAG	1600

ID AAI66367 standard; cDNA; 1712 BP.  
 XX AAI66367;  
 AC  
 XX  
 XX  
 DT 29-JAN-2002 (first entry)  
 XX  
 XX Nicotiana tabacum phytoene synthase coding sequence #2.  
 DE  
 XX Phytoene synthase; zeta carotene desaturase; herbicide; transgenic plant;  
 KW plant growth regulator; herbicidal; tobacco; ss.  
 XX  
 XX Nicotiana tabacum.  
 OS  
 XX  
 XX  
 XX Key Location/Qualifiers  
 FH 333..1565  
 FT /\*tag= a  
 FT /product= "phytoene synthase"  
 FT  
 XX  
 XX DE10022362-A1.  
 XX  
 XX PD 15-NOV-2001.  
 XX  
 XX PF 08-MAY-2000; 2000DB-1022362.  
 XX  
 XX PR 08-MAY-2000; 2000DE-1022362.  
 XX  
 XX PA (FARF ) BAYER AG.  
 XX  
 XX PI Busch M, Hain R;  
 XX  
 XX DR WPI; 2002-027336/04.  
 XX  
 XX DR P-PSDB; AAMS1842.  
 XX  
 XX PT New nucleic acid encoding tobacco zeta-carotene desaturase, useful for  
 PT screening compounds with herbicidal activity -  
 XX  
 XX PS Claim 14; Page 21-26; 44pp; German.  
 XX  
 XX CC The present invention provides the protein and coding sequences of  
 CC phytoene synthases and zeta-carotene desaturase from Nicotiana tabacum.  
 CC The sequences can be used to identify compounds capable of altering the  
 CC expression of these genes, which are therefore useful as plant growth  
 CC regulators and herbicides. They can also be used to produce transgenic  
 CC plants. The present sequence is the coding sequence of a tobacco  
 CC phytoene synthase.  
 XX  
 XX SQ Sequence 1712 BP; 540 A; 265 C; 421 G; 475 T; 11 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 1,616-183 Length: 1712  
 Score: 1801.00 Matches: 357  
 Percent Similarity: 87.93% Conservative: 29  
 Best Local Similarity: 81.32% Mismatches: 24  
 Query Match: 79.34% Indels: 29  
 DB: 24 Gaps: 4  
 US-09-847-081b-2 (1-440) x AAI66367 (1-1712)  
 QY 3 MetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThrGly 22  
 DB 333 AGTCTGTTGCTTGTATGCGTGTTCACCT---TGTGAAGTCTCAATGGGACAGCA 389  
 QY 23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42  
 DB 390 TTCTTGATTCAGTCCGGAGGGAACCGGGTTTGTGATTCGCG-----AGG 437  
 QY 43 AspArgAsnLeuMetTrpAsnGlyArgGlyLeuGlyGlyArgGluArgTrpAsnPhe 62  
 DB 438 CATGCAATTACTGTGCAATGAGAAACAGACAGAGCTGTGAAACAAAGGTGCAATTTT 497  
 QY 63 GlySerLeuIleAlaAspProArgTrpSerCysLeuGlyGlySerArgThrGluLysGly 82  
 DB 498 GGT----- 500

QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyLysMet---ThrVal 101  
 DB 501 -----TCTGTAAAGTCTCTATGCTTACACCGGCGGAGAAATGGCGACAGT 551  
 QY 102 SerSerGluLysValTyrAspValValLeuLysGlnAlaAlaLeuValLysArgGln 121  
 DB 552 ACATCAGAACAGATGGTTATGATGCTGTTTAAACAGACGCTTGTGTAAGGCGAG 611  
 QY 122 LeuArgSerThrAspAspLeuGluValLysProAspIleValValProGlyLysnLeuGly 141  
 DB 612 TTGAATCTGCTGAATGATTTAGAACTGAAGTGAAGCCGAGATCCCTCCCGGAAATTGAC 671  
 QY 142 LeuLeuSerGluValTyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThrPhe 161  
 DB 672 TTGTTGAGTGAAGCATATGATGCTGTGTAAGTATGTGCAAGTATGCAAGCATTT 731  
 QY 162 TyrLeuGlyThrLysLeuMetThrProGluLysArgArgAlaIleTrpAlaIleTyrVal 181  
 DB 732 TACTTGGAAACCATGTYATATGACTCCAGAGAGAAAGGGCTATTGGGCAATATATGTG 791  
 QY 182 TrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGln 201  
 DB 792 TGGTCAGAGAAACAGATGAACCTTGTTGATGGCCCAACGATCATATTAACCCCA 851  
 QY 202 AlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMet 221  
 DB 852 GCCTTAGATAGTGGGAAAGACCGGCTTGAAGATGTTTCAGCGGCGACCATTTGATATG 911  
 QY 222 LeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArg 241  
 DB 912 CTCGATGCGCTGTTGTCGATCTGTTCCAAAGTTCCAGTTGATATTCAGCGTTCCGA 971  
 QY 242 AspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTrpLysThrPheAspGlu 261  
 DB 972 GATATGATTGAAGGAATGCTATGATGAGTGAAGATCAAGATATATGAACCTTATATGAG 1031  
 QY 262 LeuTyrLeuTyrCysTrpTyrValAlaGlyThrValGlyLeuMetSerValProValMet 281  
 DB 1032 CTTTACCTCTATGTTATTTACTCTGCTGATGAGTGGTGTATGAGGTTCCAATATGAG 1091  
 QY 282 GlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeu 301  
 DB 1092 GGTATTTGACCTGATTTCAAGGCAACACAGAGACGTATATTAAGCGAGCTTTGGCTTTA 1151  
 QY 302 GlyLeuAlaAsnGluLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGly 321  
 DB 1152 GGAATCGCAAACTCAACTAAGCAACATCTCAGAGATGTTGAGAGATGCCAGAGAGGA 1211  
 QY 322 ArgValTyrLeuProGluLysAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePhe 341  
 DB 1212 AGAGTCTACTTACCTCAAGATGATTAATGACAGCGAGCTCTTCGACGATACATATTT 1271  
 QY 342 AlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArgAlaArg 361  
 DB 1272 GCTGGAAGAAAGTACCTGATTAAGTGAAGAGCTTTATGAAGAACCAATCCAGAGGGCACA 1331  
 QY 362 LysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpPro 381  
 DB 1332 AGTCTCTTCAATGAGACAGAGAGAGAGATTCACAACTGACGCTCAGACGATGCGCT 1391  
 QY 382 ValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAspTyr 401  
 DB 1392 GATGGGCAATCTTCTGTGTGACCGCAAAATACGACAGAAATGAAGCCATATATGAC 1451  
 QY 402 AsnAsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIle 421  
 DB 1452 AACCACTTCACAAAGAGGCTTATGAGCAAAACCAAGAACTAATTTCTTACTACTAT 1511  
 QY 422 AlaTyrAlaLysSerLeuValProProAsnArgThr-SerSerProLeuAlaLys 439  
 DB 1512 GCTTATGCAAAATCTTGTGCCCCCTTACAGAACTTGTTCACCTTACGTAAG 1566

RESULT 5  
AAV03881  
ID AAV03881 standard; cDNA; 1814 BP.  
XX  
XX AAV03881;  
XX  
XX 29-APR-1998 (first entry)  
XX  
XX Phytoene synthase coding sequence from N. tabacum.  
DE  
XX Phytoene synthase; transgenic plant; enhanced carotenoid synthesis;  
KM ultra violet absorber; food colour; ss.  
XX  
XX Nicotiana tabacum.  
XX  
XX Key Location/Qualifiers  
FT CDS 363..1595  
FT /tag= a  
XX  
XX US5705624-A.  
XX  
XX 06-JAN-1998.  
XX  
XX 27-DEC-1995; 95US-0579667.  
XX  
XX 27-DEC-1995; 95US-0579667.  
XX  
XX 27-DEC-1995; 95US-0579667.  
XX  
XX (DELL/) DELLA-CIOPPA G R.  
PA (FITZ/) FITZMAURICE W P.  
PA (GRIL/) GRILL L K.  
PA (HELL/) HELLMANN G M.  
PA (KUMA/) KUMAGAI M H.  
PI Della-Cioppa GR, Fitzmaurice WP, Grill LK, Hellmann GM,  
PI Kumagai M;  
XX  
XX WP1; 1998-086196/08.  
XX  
XX P-PSDB; AAM41060.  
XX  
XX DNA encoding tobacco phytoene synthase polypeptides - useful for  
XX producing recombinant polypeptides or transgenic plants  
XX  
XX Claim 1; Column 33-36; 25pp; English.  
XX  
XX This sequence encodes the phytoene synthetase from Nicotiana tabacum.  
XX The phytoene synthetase coding sequence represents a cDNA of the  
XX invention. The isolated nucleic acid molecules are used for producing  
XX recombinant polypeptides or transgenic plants with enhanced ability to  
XX synthesise carotenoids. Phytoene has been used as a ultra violet absorber  
XX and other carotenoids have been used as food colours, animal feeds and in  
XX the pharmaceutical and cosmetics industries.  
XX  
XX Sequence 1814 BP; 577 A; 281 C; 441 G; 515 T; 0 other;  
SQ  
Alignment Scores:  
Pred. No.: 2,886-183 Length: 1814  
Score: 1799.00 Matches: 356  
Percent Similarity: 87.93% Conservative: 30  
Best Local Similarity: 81.09% Mismatches: 24  
Query Match: 79.25% Indels: 29  
DB: 19 Gaps: 4  
US-09-847-081B-2 (1-440) x AAV03881 (1-1814)  
QY 3 MetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThrGly 22  
Db 363 ATGCTCTGTTGCTTGTATGCGTTGTTTCACCT--TGTGAAGTCTCAATGGGACAGCA 419  
QY 23 LeuLeuAspSerValAlaGluGlyValAsnArgValPheValSerSerArgPheLeuAlaArg 42  
Db 420 TTCCTTGAGTTCAGTCCGGAGGGAACCGGGTTTTCATTGCTCG-----AGG 467  
QY 43 AspArgAsnLeuMetTrpAsnGlyArgGlyLeuGlyGlyArgGlnArgTrpAsnDhe 62

Db 468 CATAGCAATTTAGTGTCCAAATGAGAGAAACAGAGAGGTGTGAACAAAGGTGGAATTTT 527  
QY 63 GlySerLeuLeuAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluLeuGly 82  
Db 528 GGT----- 530  
QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGlyMet---ThrVal 101  
Db 531 -----TCTGTAAAGTCTGCTATGCTGTGCTTACACCGCGGGAAGAAATGGCAGCATG 581  
QY 102 SerSerGluLeuValAlaThrAspValValLeuLeuGlnAlaAlaLeuValValArgGln 121  
Db 582 ACATCACAACAGAGGTGTATGATGTGTTTAAACACAGACGCTTATGAGAACAGCAG 641  
QY 122 LeuArgSerThrAspAspLeuGluValValProAspValLeuValProGlyAsnLeuGly 141  
Db 642 TTGAGATCTGCTGATGATTTAGAGTGAAGCCGAGATCCCTCTCCCGGAATTTTAC 701  
QY 142 LeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLeuThrPhe 161  
Db 702 TTGTTAGTGAAGCATATGATAGGTATGAGGTATGTGACAGGTATGCAGAAAGCATTT 761  
QY 162 TyrLeuGlyThrLysLeuMetThrProGluArgArgAlaLeuPalaLeuTyrVal 181  
Db 762 TACTTAGAACCAGCTAATGACTCCAGAGAGAAAGGCGCTATTGGCAGATATATGTG 821  
QY 182 TrpCysArgArgThrAspGluLeuValAspGluProAsnAlaSerHisileThrProGln 201  
Db 822 TGGTCAGAGAAACAGATGAACTGTGTATGAGGCCCAACGATCATATGATACCCCA 881  
QY 202 AlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMet 221  
Db 882 GCCTTATATAGTGTGGAGAACCGGCTTGAAAGATGTTTCAGCGGCGACCATTTGATATG 941  
QY 222 LeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArg 241  
Db 942 CTCGATGCTGTTGTTCGATGACTGTTCACAGCTTCAGTTGATATTCAGCCCTTCAGA 1001  
QY 242 AspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysThrPheAspGlu 261  
Db 1002 GATATGATGAGAGAAAGCGATGAGCTTGAGAGAGTGAAGATATGAACTTTGATGAG 1061  
QY 262 LeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMet 281  
Db 1062 CTTTACTCTATTTCTTATAGTTGCTGTGACGTTGGTTGATGATGATTTCCAAATATG 1121  
QY 282 GlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeu 301  
Db 1122 GGTATTCACCTGATTCAAAGGCAACAGAGAGCTATATATATGACGCTTTGGCTTTA 1181  
QY 302 GlyLeuAlaAsnGluLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGly 321  
Db 1182 GGAATCGGAATCACTACGACATACCTCAGGATGTGGAGAAATGCCAGAAAGGAGA 1241  
QY 322 ArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePhe 341  
Db 1242 AGAGTCTACTTACTCTCAAGATGATTAAGACAGGCAAGCTCTTCGACATGACATATTT 1301  
QY 342 AlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArgAlaArg 361  
Db 1302 GCTGGAAGAAGTGAAGTGAAGTGAAGCTTTATGAGAAAGCAAAATCCAGAGGGCAAGA 1361  
QY 362 LysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpPro 381  
Db 1362 AAGTCTTTCATGAGGAGAGAGGAGGATTAACAACTGAGCTAGAGAGATGGCCT 1421  
QY 382 ValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGlnAlaAspTyr 401  
Db 1422 GTATGGGACATCTTGGCTGTGTACCGCCAAATCTGACAGAGTTGAAGCAATGATCTAC 1481  
QY 402 AsnAsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIle 421



Dd	1482 AACAACTTCACAAAAGAGACTTATGTGAGCAACAACAAAAGACTAATTCCCTACTATTT
Oy	442 AlaTYraIalabvSSerleuValProPobanaqfthr-SerSerProleuAlalyS 439
Dd	1542 GCTTATGCAMAAATCTCTGTGGCCCCCTACAGAACTCTGTGCACCTTAGCTTAAG 1596
RESULT 6	
ID	AAV03878 standard; cDNA; 1795 BP.
XX	AAV03878;
AC	
XX	
DT	29-APR-1998 (first entry)
XX	
DE	Phytoene synthase coding sequence from N. benthamiana.
XX	
KW	Phytoene synthase; transgenic plant; enhanced carotenoid synthesis;
XX	
XX	ultra violet absorber; food colour; ss.
OS	Nicotiana benthamiana.
XX	
FH	Key Location/Qualifiers
FT	CDS 361..1594
FT	/+tag= a
XX	
FN	US5705624-A.
PD	
XX	06-JAN-1998.
PF	27-DEC-1995; 95US-0579667.
XX	
PR	27-DEC-1995; 95US-0579667.
XX	
PA	(DBLJ/) DELLA-CIOPPA G R.
PA	(FITZ/) FITZMAURICE W P.
PA	(GRIL/) GRILL L K.
PA	(HELL/) HELLMANN G M.
PA	(KUMA/) KUMAGAI M H.
XX	
Pi	della-Cioppa GR, Fitzmaurice WP, Grill LK, Hellmann GM;
PI	Kumagai MH;
XX	
DR	WP1_1998-086196/08.
XX	
DR	P-BSDb; AAW41057.
XX	
PT	DNA encoding tobacco phytoene synthase polypeptides - useful for
PT	producing recombinant polypeptides or transgenic plants
PS	Claim 1; Column 15-20; 25pp; English.
XX	
CC	This sequence encodes the phytoene synthetase from Nicotiana benthamiana.
CC	The phytoene synthetase coding sequence represents a cDNA of the
CC	invention. The isolated nucleic acid molecules are used for producing
CC	recombinant polypeptides or transgenic plants with enhanced ability to
CC	synthesize carotenoids. Phytoene has been used as a ultra violet absorber
CC	and other carotenoids have been used as food colours, animal feeds and in
CC	the pharmaceutical and cosmetics industries.
XX	
SQ	Sequence 1795 BP; 577 A; 271 C; 434 G; 513 T; 0 other;
Alignment Scores:	
Pred. No.:	1.39e-180 Length: 1795
Score:	1774.00 Matches: 351
Percent Similarity:	87.07% Conserves: 33
Best Local Similarity:	79.59% Mismatches: 28
Query Match:	78.15% Indels: 29
DB:	19 Gaps: 4
US-09-847-081B-2 (1-440) x AAV03878 (1-1795)	
OY	1 MetSerMetSerValAlaleuLeuTrpValValSerProThrSerGlValSerangly 20
Dd	356 CTCAGAAATCTCGTGCCCTGTGTATGGGTGTTCACCT--TGTAAGGCTCAAAATGG 412

QY	21	ThrGlyLeuLeuAAsSerValArgGluGlySerNAValPheValSerSerTrpPheLeu	40
Db	413	ACAGAGATTCCTGGATTCAATTCCGAGGAAACCGGGTTTTGATTGGCCAGTAT---	469
QY	41	AlaArgAspArgAsnLeuMetTrpAsnGlyArgIleLysLysGlyValArgGlnArgTrp	60
Db	470	-----AGAAATTATGTGTGCATAGAGGAACAAGAGCTGGGAAACAAGGTGG	520
QY	61	AsnPhgGlySerLeuIleAlaAspProArgTrpSerCysLeuGlyGlySerArgTrpGlu	80
Db	521	AAATTGTG-----	529
QY	81	LysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMet---	99
Db	530	-----TCTGTAAAGCTCTGCTATGCTGGCTACACCGCGGAGAAATGGCG	574
QY	100	ThrValSerSerGlyLysLysValTyrAspValValIleLysGlnAlaAlaLeuValLys	119
Db	575	ACGAAGCATCAGAACATGGTTATATGATGGTTTAAACAAGACACTTATGAGGA	634
QY	120	ArgGlnLeuArgSerThrAspAspLeuGlnValLysProAspIleValValProGlnAsn	139
Db	635	AGGCAGTTGAGATCTACTGATGATTTAGAGTGAAGCCGGAATCCCTCCAGGCAAT	694
QY	140	LeuGlyLeuLeuSerGlnAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLys	159
Db	695	TTGAGCTTGTGAGTGAAGCATATGATAGTGTACTGAAAGTATGTGCAAGATATCGAAG	754
QY	160	ThrPheTyrIleuGlyThrLysLeuMetThrProGlnArgArgAlaIleTrpAlaIle	179
Db	755	ACATTATCTTAGGAACCATGCTAAATGACTCCAGAGAGAAGAGGACTATTTGGCGAAAT	814
QY	180	TyrValTyrCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThr	199
Db	815	TATGATGTGTCAGAGGAACAGATGAATCTTGTTGATGGCCGAATGCATATTTAC	874
QY	200	ProGlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPhe	219
Db	875	CCACAAAGCTTAGAAGATGGGAAACCTGCTGGAAGATGTTTTCAAGGGGGCGCATTT	934
QY	220	AspMetLeuAspAlaAlaLeuSerSerPheValSerArgPheProValAspIleGlnPro	239
Db	935	CATATGCTCGATGCTGCTTTGTCCGATCTGTTCCAAAGTTTCCAGTTGATTTACGCA	994
QY	240	PheArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysThrPhe	259
Db	995	TTCAAGATATGATGATTGAAGATAGCTGTATGACTTGAGGAAGTCAAGATACAGAACTTT	105
QY	260	AspGluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValPro	279
Db	1055	GATGAGCTATACCTATATTTGTTATTAACGTGCTGTACGGTGGGTGATGAGTGTCCA	111
QY	280	ValMetGlyIleAlaProGluSerLysAlaThrTrpGluSerValTyrAsnAlaAlaLeu	299
Db	1115	ATTATGGGTATTTGCCCTGATTTCAAAAGCACAAACAGAAATCTATATATATGACGTTTG	117
QY	300	AlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArg	319
Db	1175	GCTTTGGGTATATAGCAATCAACTCAACAACTACAGAGATGTCGAGAAATATGCCACA	123
QY	320	ArgGlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAsp	339
Db	1235	AGAGGAAGGCTTACTTACCTCAATGATTAATTAAGCATGCAAGGTCTCTCCAGATGAC	129
QY	340	IlePheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArg	359
Db	1295	ATATTCCGTGGAATAAGTACGATTAATGAGAAACCTTATGAAACCAATATCCAGAG	135
QY	360	AlaArgLysPhePheAspGluSerGluLysGlyValIleTrpGluLeuAspSerAlaSerArg	379
Db	1355	GCAGAAAGTTCTTTGATGAGGCAAGAGAGATGACACACTAGAGCTCACTAGTATA	141

QY 360 TrpProValIleuThrAlaIleuLeuTyArgLysIleLeuAspGluIleGluAlaAsn 399  
 |||||  
 Db 1415 TGGCTGTATGGGCTTCTTGTGTACCGCCAGATACTGACGAGTGAAGCCAT 1474  
 |||||  
 QY 400 AspTyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeu 419  
 |||||  
 Db 1475 GACTACCAACACTTACAAAGAGACTTATGACCAACATCAAAAGACTTATTCCTTA 1534  
 |||||  
 QY 420 ProIleAlaTyrAlaLysSerLeuValProProAsnArgThr-SerSerProLeuAlaLys 439  
 |||||  
 Db 1535 CCTATTGCTATGCAAAATCTCTTGTGCCCCCAAGAACTCTTGTCTTCTTAAGCTTA 1594  
 |||||  
 QY 439 s 439  
 |||||  
 Db 1595 G 1595  
 |||||  
 RESULT 7  
 AAV03879  
 ID AAV03879 standard; cDNA, 1316 BP.  
 AC AAV03879;  
 XX  
 DT 29-APR-1998 (first entry)  
 XX  
 DE Phytoene synthase coding sequence from N. benthamiana.  
 XX  
 KW Phytoene synthase; transgenic plant; enhanced carotenoid synthesis;  
 XX ultra violet absorber; food colour; ss.  
 OS Nicotiana benthamiana.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1242  
 FT /\*tag= a  
 XX  
 XX US5705624-A.  
 XX  
 PD 06-JAN-1998.  
 XX  
 PE 27-DEC-1995; 95US-0579667.  
 XX  
 PR 27-DEC-1995; 95US-0579667.  
 XX  
 PA (DBL//) DELLA-CIOPA G R.  
 PA (FITZ//) FITZMAURICE W P.  
 PA (GRILL//) GRILL L K.  
 PA (HEIL//) HEILMANN G M.  
 PA (KUMA//) KUMAGAI M H.  
 XX  
 PI Della-Ciopa GR, Fitzmaurice WP, Grill LK, Hellmann GM,  
 PI Kumagai MH;  
 XX  
 DR WPI, 1998-086196/08.  
 DR P-PSDB; AAW41058.  
 XX  
 PT DNA encoding tobacco phytoene synthase polypeptides - useful for  
 PT producing recombinant polypeptides or transgenic plants  
 XX  
 PS Claim 1; Column 21-26; 25pp; English.  
 XX  
 XX This sequence encodes the phytoene synthetase from Nicotiana benthamiana.  
 CC The phytoene synthetase coding sequence represents a cDNA of the  
 CC invention. The isolated nucleic acid molecules are used for producing  
 CC recombinant polypeptides or transgenic plants with enhanced ability to  
 CC synthesise carotenoids. Phytoene has been used as a ultra violet absorber  
 CC and other carotenoids have been used as food colours, animal feeds and in  
 CC the pharmaceutical and cosmetics industries.  
 CC  
 XX Sequence 1316 BP; 403 A; 221 C; 342 G; 350 T; 0 other;  
 SQ  
 Alignment Scores: 1.73e-179 Length: 1316  
 Pred. No.: 1762.00 Matches: 347  
 Score:

Percent Similarity: 87.04% Conservative: 29  
 Best Local Similarity: 80.32% Mismatches: 28  
 Query Match: 77.62% Indels: 28  
 DB: 19 Gaps: 4  
 US-09-847-081b-2 (1-440) x AAV03879 (1-1316)  
 QY 3 MetSerValAlaLeuLeuThrPValValSerProThrSerGluValSerAsnGlyThrGly 22  
 |||||  
 Db 1 ATGCTGTGCTTGTATGAGTTGTTCACT---TGTAGGCTCAAAATGGACAGGA 57  
 |||||  
 QY 23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42  
 |||||  
 Db 58 TTCTTGATTCATCCCGAGAGGAAACCGGTTTGTATGTCG-----AGG 105  
 |||||  
 QY 43 AspArgAsnLeuMetTrpAsnGlyArgGlyLeuGlyGlyValArgGlnArgTrpAsnPhe 62  
 |||||  
 Db 106 CATAGCAATTTAGTGCAATGAGAGAAACAGAGAGTGTGCAACAAATGTGAAATTT 165  
 |||||  
 QY 63 GlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluLysGly 82  
 |||||  
 Db 166 GGT----- 168  
 |||||  
 QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMet---ThrVal 101  
 |||||  
 Db 169 -----TCTGTAAGCTCGCTATGTGGCTACACCGCGGAGAAATGGCGACGATG 219  
 |||||  
 QY 102 SerSerGluLysLysValTyrAspValValLeuLysGlnAlaAlaLeuValLysArgGln 121  
 |||||  
 Db 220 ACATCAGAACAGAGTGTATGATGTGTATGAAACAACAGCTTATGTGAAGAGCGAG 279  
 |||||  
 QY 122 LeuArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeuGly 141  
 |||||  
 Db 280 TTGAGATCTACTATGATTTAGAAAGTGAAGCGGAGATCCCTCCCGGGGAATTTGAGC 339  
 |||||  
 QY 142 LeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThrPhe 161  
 |||||  
 Db 340 TTGTTGAGTGAACATGATGATGATGAGTGAAGCAAGTATGTCAGAGTGTGGAAGACATTT 399  
 |||||  
 QY 162 TyrLeuGlyThrLysLeuMetThrProGluArgArgArgAlaIleTrpAlaIleTyrVal 181  
 |||||  
 Db 400 TACTTAGAACTATGATTAATGACTCCAGAGAGAAAGGGCTATTTGGCAATATATGTA 459  
 |||||  
 QY 182 TrpCysArgArgTrpAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGln 201  
 |||||  
 Db 460 TGTGCAAGAGAACAGACGAACTTGTGATGCCCCGAATGCATATTAATTAATCCACAA 519  
 |||||  
 QY 202 AlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMet 221  
 |||||  
 Db 520 GCCTTAGATAGTGGGAAGACCGGCTGGAAGATGTTTCAGTGGCGGCATTTGACATG 579  
 |||||  
 QY 222 LeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArg 241  
 |||||  
 Db 580 CTCGATCTGCTGTTGTCGATACGTTCCAGTTCCAGTTGATATGACCGCTTGAGA 639  
 |||||  
 QY 242 AspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysTrpPheAspGlu 261  
 |||||  
 Db 640 GATATGATCGAAGAAAGCGATGACTTGAAGAAAGTCGAGATACAGAACTTTGAGAG 699  
 |||||  
 QY 262 LeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMet 281  
 |||||  
 Db 700 CTATACCTAATATGTTATTTAGTTGCTGTGATCAGTTGGTGTATGATGTTCCAAATATG 759  
 |||||  
 QY 282 GlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeu 301  
 |||||  
 Db 760 GGCATTCGACCTGATCAAAGGCAACAGAGAGTATATATATGACGCTTGGCTTGG 819  
 |||||  
 QY 302 GlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGly 321  
 |||||  
 Db 820 GGTATTCGGAATCACTACCAACATTTCTCAGAGATTCGAGAGAAATGCGAAGAGAGA 879  
 |||||  
 QY 322 ArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePhe 341  
 |||||

Db 880 AGAGTCTACTTACCTCAGATGAATTAGCAGACGAGTCTCTCCGACGATGACATATT 939  
 Qy 342 AlAGlyArgValThrAspLysTrpArgAspMetLysValGlnIleGlnArgAlaArg 361  
 Db 940 ACTGGAAAAGTACGTATTAATGGAGAACCTTTATGTAAGAACATTCAGAGGGCAGA 999  
 Qy 362 LysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpPro 381  
 Db 1000 AAGTCTTCAATGAGGACAGAGAGGAGTTACACAACTGAGCTCAGCTGACGATGGCT 1059  
 Qy 382 ValLeuThrAlaLeuLeuLeuTyraArgLysIleLeuAspGlnIleGlnAlaAspLys 401  
 Db 1060 GATGGGACATCTTGTGTGTGTAACGCCAAATACCTCGACAGATCGAACCCATGACTAC 1119  
 Qy 402 AsnAspPheThrArgArgAlaTyraValSerLysProLysValLeuLeuThrIleuProIle 421  
 Db 1120 AACAACTTACAAAGAGACCTTATGTGACCAATTCAAAGACGTAATTTCTTACTATT 1179  
 Qy 422 AlaTyraIaLysSerLeuValProProAsnArgThr 433  
 Db 1180 GCTTATGCAAAATCTCTGTGCCCCCTACAGAACT 1215

## RESULT 8

AAQ99323 standard; cDNA; 1591 BP.

AAQ99323;

13-APR-1996 (first entry)

Melon phytoene-synthase gene.

melon; phytoene-synthase; ripening; cDNA library; fruit; ME15;  
 tomato; TOM5; probe; hybridisation; polymerase chain reaction; PCR;  
 antisense; transgenic plant; crop improvement; carotenoid; vector;  
 88.

Cucumis melo.

MO9602650-A2.

01-FEB-1996.

06-JUL-1995; 95WC-GB01603.

22-SEP-1994; 94GB-0019081.

18-JUL-1994; 94GB-0014505.

(ZENEC) ZENECAL LTD.

Grierson D, John I, Karvouni Z, Taylor J, Turner A,

Watson C;

WPI, 1996-105912/11.

New isolated DNA encoding melon phytoene synthase - used to  
 transform plants to modify carotenoid content and related  
 characteristics in plant parts, partic. fruit  
 Claim 2; Page 15-16; 22pp; English.

The sequence encodes melon phytoene-synthase (ME15 gene), and is  
 almost full-length. The sequence has been isolated as a cDNA clone  
 from a ripening-related cDNA library derived from climacteric melon  
 fruit, using the tomato phytoene-synthase cDNA (TOM5) as a  
 heterologous probe. The ME15 gene 5'-end has also been isolated by  
 polymerase chain reaction and sequenced. The DNA may be used in  
 sense or antisense constructs to modify gene expression in plants.  
 The carotenoid content and related characteristics of plant parts  
 (particularly fruit) may be modified in this way.

Sequence 1591 BP; 507 A; 237 C; 395 G; 452 T; 0 other;

## Alignment Scores:

Pred. No.: 3,48e-178 Length: 1591  
 Score: 1751.00 Matches: 344  
 Percent Similarity: 86.67% Conservative: 33  
 Best Local Similarity: 79.08% Mismatches: 34  
 Query Match: 77.14% Indels: 24  
 DB: 17 Gaps: 3

US-09-847-081B-2 (1-440) x AAQ99323 (1-1591)

Qy 1 MetSerMetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAngly 20  
 Db 258 CTCAGAAATGTCGTGTGCTCTGTATGAGGTGTTCTCTCT--TGTGAGCTCTCAATGGG 314  
 Qy 21 ThrGlyLeuLeuAspSerValArgGluLysAsnArgValPheValSerSerArgPheLeu 40  
 Db 315 ACAAGCTTACAGAAATCAGTCGCGAGAGGAAACCGTTTTTGGATTCAATCG----- 365  
 Qy 41 AlaArgAspArgAsnLeuMetTrpAsnGlyArgIleLysGlyGlyArgGlnArgTrp 60  
 Db 366 ---AGCATAGGAATTTGTGTGTCCTCAATGAGCAATCAATAGAGT----- 407  
 Qy 61 AsnPheGlySerLeuIleAlaAspProArgTrpSerGlyLysGlySerArgTrpGlu 80  
 Db 408 -----GGTGAAAGCAACTAAT 425  
 Qy 81 LysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThr 100  
 Db 426 AATGACGAGAAATTTCTGTACGGCTGCTATTTTGGCTACTCCATCGAGAGAACGGACG 485  
 Qy 101 ValSerSerGluLysValTyraValValLeuLysGlnAlaAlaLeuValLysArg 120  
 Db 486 ATGACATCGAAGACAGATGCTCATGATGTGTTTGAGCAGCGACCTGTGTGAAGAG 545  
 Qy 121 GlnLeuArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeu 140  
 Db 546 CAACTGAGATCTACCAATGAGTGAAGTGAAGCGGATATACCTATTCGGGGGAATTTG 605  
 Qy 141 GlyLeuLeuSerGluAlaTyraAspArgGlyGlyValGlyAlaGluTyraAlaLysThr 160  
 Db 606 GCGTTGTGAGGAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 665  
 Qy 161 PheTyraLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrpAlaIleTyra 180  
 Db 666 TTTTAACTTAGAACAATGCTAATGACTCCCGAGAGAAAGGCTATTCGGCAATATAT 725  
 Qy 181 ValTrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrPro 200  
 Db 726 GTATGTCAGAGAAAGACATGAACTTGTGATGGCCCAACGATCATATATTAATCCCG 785  
 Qy 201 GlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAsp 220  
 Db 786 GCAGCCTTAGATGATGAGTGGGAAATAGGCTAGAAAGTGTTCATAGGGCGGCAATTTAC 845  
 Qy 221 MetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPhe 240  
 Db 846 ATGCTCGATGGCTGCTTGTCCGATACAGTTTCTAATTTCCAGTTGATATTAATCCCATTC 905  
 Qy 241 ArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyraLysThrPheAsp 260  
 Db 906 AGAGATGATGATGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 965  
 Qy 261 GlnLeuTyraLysTyraTyraValAlaGlyThrValGlyLeuMetSerValProVal 280  
 Db 966 GAACATATTCCTTATGTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1025  
 Qy 281 MetGlyIleAlaProGluSerLysValThrThrGluSerValTyraAsnAlaAlaLeuAla 300  
 Db 1026 ATGGATATCGCCCTCGAATCAAGGCAACACAGAGACGATATATTAATGCTGCTTGGCT 1085  
 Qy 301 LeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArg 320  
 Db 1086 CTGGGGATCGCAATCATTAATCACTCAAGATGTTGGAGAGATGCCGAGAGA 1145

QY 321 GlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIle 340  
 DB 1146 GGAAGAGCTACTTGGCTCCAGATGATATAGACAGCGAGGTCTATCCGATGAAGATATA 1205  
 QY 341 PheAlaGlyArgValThrAspLysTrpArgAspMetLysGlnIleGlnArgAla 360  
 DB 1206 TTTGCTGAGAGGTCACGATGAATGAGAAATCTTATGAAAGAAATCATATGAGGCA 1265  
 QY 361 ArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrp 380  
 DB 1266 AGAAAGTCTTGTGAGAGCGAGAAAGCGTGACAGAAATGAGCTCAGCTAGTAGATTTC 1325  
 QY 381 ProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAspAsp 400  
 DB 1326 CCGTATGGGCACTTGTGGTCTGTACCGCAAAATCACTAGATGAAAGCCATGAC 1385  
 QY 401 TyrAsnAspMetThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuPro 420  
 DB 1386 TACAACAACCTTCACAAAGAGAGCATATGTGAGCAATCAAGAAAGTTGATTCATTACT 1445  
 QY 421 IleAlaTyrAlaLysSerLeuValProProAsnArgThrSerSer 435  
 DB 1446 ATGCAATAGCAAAATCTTGTGCTCTCTACAAAACCTGCTCT 1490

RESULT 9  
 AAV17247  
 ID AAV17247 standard; DNA; 1239 BP.  
 AC AAV17247;  
 XX 28-MAY-1998 (first entry)  
 DT  
 XX MFOM5, modified phycoene synthase gene.  
 DE  
 XX MFOM5; phycoene synthase; chloroplast targeting sequence; enzyme; tomato;  
 KM protein expression enhancement; transgenic plant; carotenoid synthesis;  
 KW lycopene; ds.  
 XX  
 OS Lycopersicon esculentum.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1239  
 FT /\*tag= a  
 FT /note= "contains an intron"  
 FT exon 1..960  
 FT /\*tag= b  
 FT /number= 1  
 FT intron 961..990  
 FT /\*tag= c  
 FT /number= 1  
 FT exon 991..1236  
 FT /\*tag= d  
 FT  
 XX  
 XX MO9746690-A1.  
 XX  
 XX 11-DEC-1997.  
 XX  
 XX 23-MAY-1997; 97WO-GB01414.  
 XX  
 XX 07-JUN-1996; 96GB-0011981.  
 XX  
 XX (ZENNE ) ZENNECA LTD.  
 XX  
 XX Bird CR, Drake CR, Schuch WW;  
 XX  
 XX WPI; 1998-042198/04.  
 XX  
 XX P-PSDB; AAW41374.  
 XX  
 XX Enhancing gene expression without or with reduced co-suppression -  
 XX using altered DNA producing different RNA but same protein as  
 XX natural gene, useful especially in plants to allow overexpression of  
 XX a protein

XX  
 PS Claim 9; Page 15-16; 32pp; English.  
 XX  
 CC This sequence represents the modified phycoene synthase gene MFOM5, which  
 CC is also a chloroplast targeting sequence. This sequence is used in the  
 CC method of the invention for enhancing expression of a protein by an  
 CC organism, comprising inserting into its genome a nucleotide sequence  
 CC which produces different RNA on transcription to that of the gene already  
 CC present, but produces the same protein on translation. Transgenic plants  
 CC with enhanced ability to express a selected can be produced by the  
 CC method. For example, the method can be used to achieve overexpression of  
 CC a gene specifying an enzyme necessary for carotenoid synthesis in plants  
 CC (especially phycoene synthase), to enhance carotenoid expression,  
 CC e.g. overexpression of the carotenoid lycopene responsible for the red  
 CC coloration of developing tomato fruit. Protein expression is enhanced by  
 CC inserting a gene construct which is altered by maximizing the  
 CC dissimilarity of nucleotide usage whilst maintaining identity of the  
 CC encoded protein. Known methods of increasing protein production by gene  
 CC insertion sometimes result in low or no expression (co-suppression),  
 CC especially when the recombinant and endogenous gene sequences are  
 CC similar. The method allows enhanced expression whilst avoiding or  
 CC reducing co-suppression, since sequence similarity between the two genes  
 CC is sufficiently reduced.  
 CC  
 XX  
 SQ Sequence 1239 BP; 343 A; 247 C; 320 G; 329 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 137e-177 Length: 1239  
 Score: 1744.00 Matches: 343  
 Percent Similarity: 86.61% Conservative: 32  
 Best Local Similarity: 79.21% Mismatches: 34  
 Query Match: 76.83% Indels: 24  
 DB: 19 Gaps: 3

US-09-847-081B-2 (1-440) x AAV17247 (1-1239)  
 QY 3 MetSerValAlaLeuLeuTyrValValSerProThrSerGluValSerAsnGlyThrGly 22  
 DB 1 ATGAGCGTGGCACTTCTTGGTGGTGTAGCCCA---TGCGATGTGAGTAAACGGCACTTCA 57  
 QY 23 LeuLeuAspSerValAlaGlyGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42  
 DB 58 TTTATGAGAGGTGTGAGAGAGGTAAATGATTTCTTCAACAGTTCT-----CGT 105  
 QY 43 AspArgAsnLeuMetTPAsnGlyArgGlyLeuSerGlyArgGlnArgTrpAsnPhe 62  
 DB 106 CACCTTAACCTTGTATTAACGAACGTAATAACGGGGA----- 144  
 QY 63 GlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluLysGly 82  
 DB 145 -----CGAGGTAAACAGCAAAACACGCT 168  
 QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThrValSer 102  
 DB 169 AGAAAGTCTTCAGTTAGATCAGCAATCTTGCAACACCTGCGGTGAGAAACTATACT 228  
 QY 103 SerGluLysLysValTyrAspValValLeuLysGlnAlaLeuValLysArgGlnLeu 122  
 DB 229 AGCAGCAAAATGTGTACAGCGTGTACTTGTAACTGTACACTAGTTAAACGTCAAGTTA 288  
 QY 123 ArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeuGlyLeu 142  
 DB 289 CGTAGTACTAAACGAACCTTAGAGTTAAACGTGACACTTCAATACCTGAAACCTTGACTT 348  
 QY 143 LeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThrPheTyr 162  
 DB 349 CTTTCTAGGCTTACGACAGATGCGAGAGGTTTCCACAGAAATACGCTTAAACCTTCAAT 408  
 QY 163 LeuGlyThrLysLeuMetThrProGluArgArgAlaIleTyrPalaIleTyrValTyr 182  
 DB 409 TTGGGTACCAATGTGAGACCCAGAAAGCGTGTGTCAATATGCGCTATTTAGCTTTGG 468  
 QY 183 CysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGlnAla 202

```

Db      469  TGTAGCCGCTACTGACGAGTGTAGTGAACGACCTAATGCTAGTATACATAACACCCGCTGCT
Qy      203  LeuAaPrgTrrpGluThrArgLeuGluAaPrlPheSerGlyArgProPheaspMetLeu
Db      529  CTTCAGACATGGGAGAACCGCTTTGGAGAGACGTGTTTAACGAGACCTTCCATATGTTG
Qy      223  AspaLaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArgAaP
Db      589  GACGAGACACTTAGTACACTGCTGAGCAATTTCCCTGTGACATCCAACTTTTGGGAGC
Qy      243  MetIleGluGlyMetArgMetAspLeuTrrpLysSerArgTrrpLysThrPheaspGluLeu
Db      649  ATGATCGAGGGGATGAGATGATCTTCCTAAGTCTCGTTATAGAAATTTGATGATGTTG
Qy      263  TrrLeuTrrCysTrrTrrValAlaGlyThrValGlyLeuMetSerValProValMetGly
Db      709  TATTGTACTGCTACTACGTCGACGAGAACCGTGGCCCTTATGTCAGTCCCTATTCATGGA
Qy      283  IleAlaProGluSerLysAlaThrThrGluSerValTrrAsnAlaAlaLeuAlaLeuGly
Db      769  ATTGCACCAAGAGATGAAGCTACTACTGTAATCTGTTTAAACCGCAGCACTACATTAAGT
Qy      303  LeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGlyAspAlaArgArgGlyArg
Db      829  ATAGCTAACCGACTTACAAATATCTTGAGGAGCGTGGTGAGACGACGACGAGGCTCGT
Qy      323  ValTrrLeuProGluAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAla
Db      889  GTGATCTCCACAGACAGACGCTCGCTCAAGCTGAGTTGAGTGAACGAGACATTTTCGCA
Qy      343  GlyArgValThrAspLysTrrPargAsnPheMetLysGlyGlnIleGlnArgAlaArgLys
Db      949  GGTCTGTATTACAGACAGAGTGAAGATTTTCATGAAAGAACCAATCCCTCTCTCTAA
Qy      362  PhePheAspGluSerGlyLysGlyValThrGluLeuAspSerAlaSerArgTrrProVal
Db      1009  TTTTTCGAGCAAGCTGAAAGGAGAGTACTGAGCTTTCAGTGCATCAAGGTTCCAGTT
Qy      382  363  PhePheAspGluSerGlyLysGlyValThrGluLeuAspSerAlaSerArgTrrProVal
Db      1068  1009  TTTTTCGAGCAAGCTGAAAGGAGAGTACTGAGCTTTCAGTGCATCAAGGTTCCAGTT
Qy      402  363  LeuThrAlaLeuLeuLeuTrrArgLysIleLeuAspGluIleGlnAlaAsnAspTrrAsn
Db      1128  1069  TGGGCCAGCTGTGCTCTATAGAAAGATTTTGAGCAAGATGAGGCTTAACGATTATAT
Qy      422  403  AsnPheTrrArgArgAlaTrrValSerLysProLysLysLeuLeuThrLeuProIleAla
Db      1188  1129  AATTTTACTTAACGTCCTTACCTTCTTAAGACCAAAAACCTTATGCTCTTCAATCGCT
Qy      435  423  TrrAlaLysSerLeuValProProAsnArgTrrSerSer
Db      1227  1189  TACGCTAAGAGCTGTGTTCCACCACTAAGACAGCTAGC

```

```

FT      /transl_except= (pos: 1057..1059, aa: Arg)
XX      MO200009722-A2.
PN      24-FEB-2000.
XX      10-AUG-1999; 99MO-US18066.
PF      10-AUG-1998; 98US-0096111.
PR      07-JUN-1999; 99US-0137977.
XX      (MONS ) MONSANTO CO.
XX      Brown SM, Ellich TD, Heck GR, Kishore GM, Loguech EW, Logusch SJ,
PI      Piller KJ, Rao S, Ream JE,
XX      WPI: 2000-224351/19.
DR      F-PSDB; AAY84101.
XX      Obtaining transgenic plant useful for controlling seed germination and
PT      seedling growth comprises transgene comprising a sequence expressing
PT      altered levels of an essential hormone
XX      Claim 45; Page 254-255; 267pp; English.
XX      The present sequence encodes a phytoene synthase polypeptide, which
CC      is used in the method of the invention. The specification describes
CC      methods for the inhibition and control of gibberellic acid levels.
CC      Gibberellic acid levels may be inhibited or controlled by use of
CC      a chimeric expression construct expressing a RNA or protein which
CC      suppresses the gibberellin biosynthetic pathway sequence, diverts
CC      substrate from the pathway, or degrades pathway substrates or products.
CC      The method uses copoly1 diphosphate synthase, 3beta-hydroxylase,
CC      2-oxidase, phytoene synthase, C-20 oxidase, and a 2beta,3beta-hydroxylase
CC      polynucleotides to achieve this. The method is used to control seed
CC      germination and seedling growth especially to regulate gene products of
CC      gibberellin biosynthetic pathway and restoration of normal seed
CC      germination, in transgenic plants. The plants produced are gibberellin
CC      deficient, and have shortened hypocotyl and/or epicotyl phenotypes
CC      compared to normal plants.
XX      SQ      Sequence 1239 BP; 376 A; 208 C; 316 G; 339 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.: 7 756-177 Length: 1239
XX      Score: 1737.00 Matches: 342
XX      Percent Similarity: 86.37% Conservative: 32
XX      Best Local Similarity: 76.98% Mismatches: 35
XX      Query Match: 76.52% Indels: 24
XX      DB: 21 Gaps: 3
XX
XX      US-09-847-081B-2 (1-440) x AA299482 (1-1239)
Qy      3  MetSerValAlaLeuLeuTrrPargValSerProThrSerGluValSerAsnGlyThrGly
Db      1  ATGCTGTGCTGCTTATAGGTTGTTCTCTCT--TGTGACGTCTCAATGAGCAAGT
Qy      23  LeuLeuAspSerValArgGlyLysAsnArgValPheValSerSerArgPheLeuAlaArg
Db      58  TTCATGGAATCAGTCCGGAGAGGAAACCGTTTGTGATTCATCG-----AGG
Qy      43  AspaArgAsnLeuMetTrrPargLysArgLysLysGlyValArgGlnArgTrrAsnPhe
Db      106  CATAGGAATTTGTGTTCCATAGAGAAATCAATAGAGT-----
Qy      63  GlySerLeuIleAlaAspProArgTrrSerCysLeuGlyGlySerArgThrGluLysGly
Db      145  -----GGTGGAAAGCAAACTAATATGGA
Qy      83  SerTrrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGlyMetThrValSer
Db      169  CGGAATATTTCTGTAGCGTCTGCTATTTTGGCTACTCATCTGAGAAACGAGCATGACA

```

OY	103	Ge	g	u	l	y	s	l	y	s	v	a	l	t	y	r	a	s	p	a	v	a	l	e	u	s	g	i	n	a	a	l	e	v	a	l	y	r	a	g	l	e	u	122																	
		229	T	G	G	A	C	G	A	G	G	T	C	T	A	T	A	G	T	G	G	T	T	T	G	A	G	G	G	G	C	C	T	T	G	T	A	A	G	G	C	A	G	288																	
OY	123	A	r	s	e	r	t	h	r	a	s	p	a	s	p	e	l	e	u	l	e	v	a	l	e	u	s	g	i	n	a	a	l	e	v	a	l	y	r	a	g	l	e	u	142																
		289	A	G	A	T	T	C	C	A	T	A	G	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	348																		
OY	143	L	e	u	s	e	r	g	i	n	a	a	l	e	v	a	l	e	u	s	g	i	n	a	a	l	e	v	a	l	e	u	s	g	i	n	a	a	l	e	v	a	l	y	r	a	g	l	e	u	162										
		349	T	T	G	A	C	T	A	G	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	408																		
OY	163	L	e	u	g	i	n	a	a	l	e	u	s	g	i	n	a	a	l	e	v	a	l	e	u	s	g	i	n	a	a	l	e	v	a	l	e	u	s	g	i	n	a	a	l	e	v	a	l	y	r	a	g	l	e	u	182				
		409	T	T	G	A	C	T	A	G	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	468															
OY	183	C	y	a	r	a	g	r	a	t	h	r	a	s	p	a	s	p	e	l	e	u	l	e	v	a	l	e	u	s	g	i	n	a	a	l	e	v	a	l	e	u	s	g	i	n	a	a	l	e	v	a	l	y	r	a	g	l	e	u	202
		469	T	G	C	A	G	A	G	A	C	G	A	T	T	G	T	G	A	T	T	G	A	T	T	G	A	T	T	G	A	T	T	G	A	T	T	G	A	T	T	G	A	T	T	G	528														
OY	203	L	e	u	a	s	p	a	r	g	t	r	i	p	g	i	n	t	h	r	a	g	l	e	u	s	g	i	n	a	a	l	e	v	a	l	e	u	s	g	i	n	a	a	l	e	v	a	l	y	r	a	g	l	e	u	222				
		529	C	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	T	A	G	588															
OY	223	A	s	p	a	a	l	e	u	s	e	r	a	s	p	e	r	h	r	a	s	e	r	g	r	a	p	h	e	r	p	r	o	v	a	l	e	g	i	n	a	a	l	e	v	a	l	y	r	a	g	l	e	u	242						
		569	G	A	T	G	T	G	C	T	T	G	T	C	C	G	A	T	A	C	A	G	T	T	C	T	T	A	C	T	T	G	A	T	T	C	A	G	T	T	C	A	G	T	T	C	A	G	648												
OY	243	M	e	c	i	l	e	g	i	n	a	a	l	e	u	s	g	i	n	a	a	l	e	v	a	l	e	u	s	g	i	n	a	a	l	e	v	a	l	e	u	s	g	i	n	a	a	l	e	v	a	l</									

XX	22-OCT-2002	(first entry)	
DT			
XX			
DE	Tomato phytoene synthase cDNA.		
XX			
XX	Gibberellin; transgenic plant; seed germination; seedling growth;		
KW	transgenic; phytoene synthase; enzyme; GAI gene; tomato; ss.		
XX			
OS	Lycopersicon esculentum.		
XX			
EH	Key	Location/Qualifiers	
FT	CDS	1..1239	
FT		/*tag= a	
FT		/product= "Phytoene synthase"	
FT		/transl_except= (pos:1027..1029, aa:Lys)	
FT		/transl_except= (pos:1057..1059, aa:Arg)	
XX			
PN	US2002053095-A1.		
XX			
PD	02-MAY-2002.		
XX			
PF	10-AUG-1999; 99US-0371307.		
XX			
PR	10-AUG-1999; 99US-0371307.		
XX			
PA	(BROW/) BROWN S M.		
XX			
PI	Brown SM, Ellich TD, Heck GR, Kishore GM, Logusch EW, Logusch SJ;		
PI	Piller KU, Rao S, Ream JB;		
XX			
DR	WPI; 2002-489107/52.		
XX	P-PSDB; AAE24923.		
PT			
PT	Control of gibberellin levels in plants useful to avoid unfavorable		
PT	conditions in crops to increase yields, using transgenic plants having		
PT	reduced seed germination and early seedling growth then treatment to		
PT	restore these properties		
XX			
PS	Claim 45; Page 99; 155pp; English.		
XX			
CC	The invention relates to control of gibberellin (GA) levels in plants.		
CC	The method involves producing transgenic plants having a phenotype		
CC	of reduced seed germination and reduced early seedling growth, then		
CC	restoring seed germination and early seedling growth by treating		
CC	plants with an appropriate compound when conditions are favourable.		
CC	The method is useful to control seed germination and/or early seedling		
CC	growth in agricultural production so that unfavorable environmental		
CC	conditions normally reducing agronomic output can be avoided and		
CC	yields increased. Plants also demonstrate increased uniformity of		
CC	germination, emergence and seedling vigor, so increasing yields at		
CC	harvest. The method is especially useful in crop plants such as e.g.		
CC	canola, soybean, cotton, etc., and is also useful in storage and		
CC	transport of seeds to reduce premature germination which may affect		
CC	agronomic or food quality of the seeds. The present sequence is		
CC	tomato phytoene synthase cDNA. This cDNA is used in exemplification		
CC	of the invention.		
XX			
SO	Sequence 1239 BP; 376 A; 208 C; 316 G; 339 T; 0 other;		
	Alignment Scores:		
	Pred. No.: 7.76e-177	Length: 1239	
	Score: 1737.00	Matches: 342	
	Percent Similarity: 86.37%	Conservative: 32	
	Best Local Similarity: 78.98%	Mismatches: 35	
	Query Match: 76.52%	Indels: 24	
DB:	24	Gaps: 3	
	US-09-847-081B-2 (1-440) x AAD40272 (1-1239)		
QY			
	3 MeSeSeValAlaIeuLeuTPValValSerProThSerGluValSerAnGlyThrGly 22		
	1 ATGTCGTGTGGCTGTATAGGGTGTGTTCTCCG---TGTACCTCTCAATGGACAACT 57		

QY	23	LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuLysArg	42
Db	58	TTCAATGAAATACAGTCCGGAGGAAACCGTTTTTTGATCATCG-----AGG	105
QY	43	AspArgLeuLeuMetTrpAsnGlyArgIleuLysGlyArgGlnArgTrpAsnDhe	62
Db	106	CATAGGAATTTGGTGTCCAAATGAGAAATCAATAGAGT-----144	
QY	63	GlySerLeuIleAlaAspProArgTrpSerCysLeuGlyGlySerArgThrGluLysGly	82
Db	145	-----GGTGGAAAGCAAACTAATTAATGCA168	
QY	83	SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGluMetThrValSer	102
Db	169	CGAAATTTTCTGACGGTCTGCTCAATTTGGCTACTCCATCTGGAGAACGGACGTGACA228	
QY	103	SerGluLysLysValTrpAspValValLeuLysGlnAlaAlaLeuValLysArgGlnLeu	122
Db	229	TCGGAAACAGATGGTCTATGATGTGGTTTGGAGCGAGCGACCTTGCTGAAGCGAACTG288	
QY	123	ArgSerThrAspAspLeuGluValLysProAspIleValProGlyAsnLeuGlyLeu	142
Db	289	AGATCTACCAATGAGTTGAAAGTGAAGCCGATATCACTATTCGGGAAATTTGGGCTTG348	
QY	143	LeuSerGluAlaTrpAspArgCysGlyGlyValCysAlaGluTrpAlaLysThrPheTrp	162
Db	349	TTGATGTAGCACTAATGATAGGTGTGTGAAGTATGTGACAGTATGCAAAAGCGTTTAAAC408	
QY	163	LeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrpAlaIleTrpValTrp	182
Db	409	TTAGGAATTAATGCTAATGACTCCGACAGAAAGAGGCGCTATCTGGGCAATATATGTATGG468	
QY	183	CysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGlnAla	202
Db	469	TGCAGAAAGAACAGATGAACCTTGTTGATGGCCCAACGCATCATATATTACCCGGCAGCC528	
QY	203	LeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeu	222
Db	529	CTAAGTATAGGTGGAAATAGCGCTAAGAAATGTTCATATGGCGGCCATTGACATGTCTC588	
QY	223	AspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArgAsp	242
Db	589	GATGGTGTGTGTCCGATACAGTTTCTAACTTTCCAGTTGATATTCAAGCATTTCAAGAAAT648	
QY	243	MetIleGluGlyMetArgMetAspLeuTrpLysSerArgTrpLysThrPheAspGluLeu	262
Db	649	ATGATTTGAAGGAATCGTATGAGACTTGGAAGAAATGAGATTCMAAAACTTGACAGCACTA708	
QY	263	TyrIleuTrpCysTrpTrpValAlaGlyThrValGlyLeuMetSerValProValMetGly	282
Db	709	TACCTTTATTTGTATTATGATGTGCTGGTACGGTGGTGTGATGAGAGTGTCCAAATTAATGGGT768	
QY	283	IleAlaProGluSerLysValAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGly	302
Db	769	ATTCGCCCTGAATCAAAGCAACAACAAGACGCAATATTAATGCTGTGGCTCTGTGGGG828	
QY	303	LeuAlaGlnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGlyArg	322
Db	829	ATCGCAATTCATTAATCACTAATCACTACAGAGATGTTGGAGAAAGATCCAGAAAGGAGAGA888	
QY	323	ValTrpLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAla	342
Db	889	GCTTACTTGCTCCTCAAGATGAATTTAGACACAGCGACGCTTATCCGATGAAGATATATTTGCT948	
QY	343	GlyArgValIleTrpAspLysTrpArgAsnDheMetLysGlnIleGlnArgAlaArgLys	362
Db	949	GGAGGGGTCACCGATTAATGGGAATCTTTATGAAGAAACAATTCATATGGGCAAGAAAG1008	
QY	363	PhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpProVal	382
Db	1009	TTCTTTTATGAGGCGAGAAATGGCGCTGACAGAAATTTGACTCAGCTAGTATATTTCCCTGTA1066	
QY	383	LeuThrAlaLeuLeuTrpArgLysIleLeuAspGluIleGluAlaAsnAspTrpAsn	402

Dd		1069	TGGGCACTTGTGGTTCTTGTAACCGCAAATACTAGATGAGATTGAAGCCAAATGACTAAC	1127
Oy		403	AsnphethrArgrAlaIatYrValSerlySProlYSLeuleuThrLeuProileala	422
Dd		1129	AACCTGCACAAGAGACATATGTGAGCAAAATCAAGAAGTTGATTCATTACTATTGCA	1187
Oy		423	TyrAlalysSerleuValProProAamArgHrTherSerS	435
Dd		1189	TATGCAAAATCTCTGTGGCTCTCTCACAAAAACTGCCTCT	1227
<b>RESULT 12</b>				
ID	AAQ12495			
ID	AAQ12495 standard; cDNA; 1646 BP.			
XX	AAQ12495;			
XX				
AC				
DT	25-MAR-2003 (updated)			
DT	18-SEP-1991 (first entry)			
XX				
DE	Tomato fruit ripening related gene PTOM5.			
XX				
KW	ripening; lycopene; transgenic tomato; ss.			
XX				
OS	Lycopersicon esculentum var. Alisa Craig.			
XX				
FH	Key Location/Qualifiers			
FT	CDS 201..1436			
FT	/*tag= a			
XX				
PN	M09109128-A.			
XX				
PD	27-JUN-1991.			
XX				
PF	10-DEC-1990; 90MO-GB01924.			
XX				
PR	13-DEC-1989; 89GB-0028179.			
XX				
PA	(ICIL ) IMPERIAL CHEM IND PLC.			
PI	Bird CR, Grierson D, Schuch W;			
DR	WPJ; 1991-208154/28.			
XX				
PT	DNA construct to modify synthesis of plant carotenoid(s) -			
PT	comprises sequence homologous to gene of clone PTOM5 preceded by			
PT	plant promoter			
XX				
PS	Disclosure; Fig 1; 35pp; English.			
XX				
CC	Clone PTOM5 was derived from a cDNA library isolated from ripe			
CC	tomato RNA (Slater et al., Plant Molecular Biology 5, 137-147,			
CC	1985). The protein it encodes is estimated to have mol. wt. ca.			
CC	48kD. PTOM5 is expressed in ripening fruit. Strongest expression is			
CC	at the full orange stage of ripening; no expression is detected in			
CC	green fruit. See also AAQ12494.			
CC	(Updated on 25-MAR-2003 to correct PA field.)			
XX				
SO	Sequence 1646 BP; 529 A; 249 C; 387 G; 481 T; 0 other;			
<b>Alignment Scores:</b>				
Pred. No.:	8.5e-176	Length:	1646	
Score:	1729.00	Matches:	345	
Percent Similarity:	85.68%	Conservative:	32	
Best Local Similarity:	78.41%	Mismatches:	36	
Query Match:	76.17%	Indels:	27	
DB:	12	Gaps:	4	
US-09-847-081B-2 (1-440) x AAQ12495 (1-1646)				
Oy	1 MesServerSeValAlaleuLeuTrValValSerProThrSerGIuValSerAnclY	20		
DB	195 CTAGAAATGTCGTGGCTCTGTATGAGGTGTGTTCTCTCT--TGTGAGCTTCAAAAGGG	251		

```

Oy 21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40
Db 222 ACNAGTTTCATGAGATCACTCCGGAGGGAACCGTTTTCATGATTCATCG----- 302
Oy 41 AlaArgAspArgAsnLeuMetTrpAsnGlyArgGlyLeuGlyGlyArgGlnArgTrp 60
Db 303 ---AGGCAATGAGAAATTTGGTGTCCAAATGAGAAATCAATGAGAGT----- 344
Oy 61 AsnPheGlySerLeuLeuLeuAspProArgTrpSerCysLeuGlyGlySerArgTrpGlu 80
Db 345 -----GGTGGAAAGCAAACTAT 362
Oy 81 LysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThr 100
Db 363 AATGACGGAATTTTCTGTACGCTGTCTATTTGGCTACTCTCCATCTGAGAAAGGACG 422
Oy 101 ValSerSerGluLeuLeuValTrpAspValValLeuGlyGlnAlaLeuValLysArg 120
Db 423 ATGACATCGAAGACATGGTCTATGATGTGTTTGAGGACGACGCTTGTTGAAAGAG 482
Oy 121 GlnLeuArgSerThrAspAspLeuGluValLysProAspLileValProGlyAsnLeu 140
Db 483 CAATGAGATCTAACCAATGAGTTAGAGAGCCGAGATACCTATTCGGGGAATTTG 542
Oy 141 GlyLeuLeuSerGluAlaTyrAspArgCysGlyGlyValCysAlaGlyTyrAlaLysThr 160
Db 543 GCGTTGTTAGAGGAACATATAGTGTGTGTGTAGAGTGTGTGAGAGTATCCAAAGACG 602
Oy 161 PheTyrLeuGlyTyrLeuLeuMetThrProGluArgArgAlaLileTrpAlaLileTyr 180
Db 603 TTTAAGTTAGGAACATGCTATGACTCCCGAGAGAAAGGCTATCTGTCATATAT 662
Oy 181 ValTrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisLileThrPro 200
Db 663 GTATGTGTGAGAAAGACAGATGACTGTGTGATGAGCCCAAGCAATATATATACCCG 722
Oy 201 GlnAlaLeuAspArgTrpGluTrpArgLeuGluValAspLilePheSerGlyArgProPheAsp 220
Db 723 GCAGCTTATGATGATGAGGAATAGGCTAGAGAGTGTTCATAGGGCGGCATTATGAC 782
Oy 221 MetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspLileGlnProPhe 240
Db 783 ATGCTCGATGTGTGCTTTGTCCGATACAGTTTCTAATCTTCAGATATTCAGCAATTC 842
Oy 241 ArgAspMetLileGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysThrPheAsp 260
Db 843 AGAGATATGATTTGAAAGATGCGTATGACCTTGAGAAATCGAGATACAAATAAATTCGAC 902
Oy 261 GlnLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProVal 280
Db 903 GAACTATACCTTATTTATTTATTTATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 962
Oy 281 MetGlyLileAlaProGluSerLysAlaThrThrArgLysSerValTyrAsnAlaAlaLeuAla 300
Db 963 ATGGGTATCGCCCTCGAATCAAAAGCAACAAGAGAGGATATATATATGCTGTGCT 1022
Oy 301 LeuGlyLeuAlaAsnGlnLeuThrAsnLileLeuArgAspValGlyGluAspAlaArgArg 320
Db 1023 CTGGGAGATCGCAAACTAATTAATCACTACATACAGATGTGTGGAAGAGATGCCGAAGA 1082
Oy 321 GlyArgValTyrLeuProGluAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspLile 340
Db 1083 GGAAGAGTCTACTCTCCATGATTAATTTAGCAGGAGGAGTCTATCCGATGAGATATTA 1142
Oy 341 PheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysValGlnLileGlnArgAla 360
Db 1143 TTTGCTGGAAGGCTGACCGATTAATGGAATCTTTATGAGAAGAAACAATCATATGAGGCA 1202
Oy 361 ArgLysPhePheAspGluSerGlyLysGlyValThrGluLeuAspSerAlaSerArgTrp 380
Db 1203 AGAAGATCTTTGATGAGGACAGAAAGGCGTGACAGAAATGACCTCAGTATGATGTTTC 1262

```

```

Oy 381 ProValLeuThrAlaLeuLeuTyrArgLysLileLeuAspGluLileGluAlaAsnAsp 400
Db 1263 CTTGTATGGGATCTTTGGTCTGTACCGCAAAATATGATGATGATGAAAGCAATGAC 1322
Oy 401 TyrAsnAsnPheThrArgAlaGlyAlaTyrValSerLysProLys-LysLeuLeuThrLeuPr 420
Db 1323 TACAACAACCTTCACAAAGAGCATATGTGAGCAAAATCAAGCAAGTGTGATTCATTAAC 1382
Oy 420 GllAlaTyrAlaLysSerLeuValProProAsnArgThrSerSerProLeuAlaLys 439
Db 1383 TATTGATATGCAAAATCTTGTGTCTCTCT-----ACAAACTGCTCTCTTCAAA 1434

RESULT 13
AAVI6951
ID AAVI6951 standard; cDNA to mRNA; 2868 BP.
XX
AC AAVI6951;
XX
DE 06-JUL-1998 (first entry)
XX
DE Nucleic acid encoding phycoene synthase 4.
XX
KM Phycoene synthase; breeding; variable flower colour; ds.
XX
OS Gentiana lutea.
XX
FH Key location/Qualifiers
FT CDS 400..1689
/*tag= a
FT
PV JP10084966-A.
XX
PD 07-APR-1998.
XX
PF 17-SEP-1996; 96JP-0245107.
XX
PR 17-SEP-1996; 96JP-0245107.
XX
PA (IWATE-) IWATE KEN.
XX
WP1: 1998-264853/24.
DR P-PSDB; AAW6964.
XX
PT Phytoene synthase gene - useful for breeding plant of variable
PT flower colour
PS Claim 4; Pages 12-14; 15pp; Japanese.
XX
CC The present sequence encodes phycoene synthase 4. It was isolated from
CC a cDNA library prepared from mRNA extracted from the petals of Gentiana
CC lutea. The nucleic acid sequence was amplified from the library using
CC PCR primers AAVI6952-53. The phycoene synthase gene is useful for
CC breeding plants with variable flower colours.
XX
SQ Sequence 2868 BP; 887 A; 499 C; 648 G; 833 T; 1 other;

Alignment Scores:
Pred. No.: 4.98e-161 Length: 2868
Score: 1595.00 Matches: 313
Percent Similarity: 83.68% Conservative: 46
Best Local Similarity: 72.96% Mismatches: 60
Query Match: 70.26% Indels: 10
DB: 19 Gaps: 3

US-09-847-081b-2 (1-440) x AAVI6951 (1-2868)
Oy 1 MetSerMetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGly 20
Db 394 GTTAACTATGCTATTTTGTACGCTATGGCTTTCCGCGAGTTCTGAAGTTTGAGAGGCG 453
Oy 21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40
Db 454 AATGTTTCTTGTGAGCGCAATTCGAAAGT-----TACCAT 489

```



```

QY 41 AlaArgAspArgAsnLeuMetTrpAsnGlyArgGlyLeuLysGlyValArgGlnArgTrp 60
DB 490 TTTTCGGATATAAAGTTTAAATGTCATGAGAGATTAGAGAAAGTACAGCAACCAAGCGCT 549
QY 61 AsnPhenGlySerLeuLeuLeuAspProArgTrpSerCysLeuGlyGlySerArgTrpGlu 80
DB 550 AGATCAAGTTATGGGGTTGGAGATTGAGTTCAATTTGCTTGGAGAGCTGATGATGAG 609
QY 81 ---LysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGlnMet 99
DB 610 ACCCGGGAGAGAGATTATCGGTATCTCTCAGTATTATGCTACCCGGAGAGAGAAAG 669
QY 100 ThrValSerSerGlyLysValLysValLysValLysValLysValLysValLysValLys 119
DB 670 ACGATGACATCAGAGCAAAAGTTTATGATGCTGCTTTAAAGACACACCTTGATTAAT 729
QY 120 ArgGlnLeuArgSerThrAspAspLeuGlyValLysProAspLysValLysValProGlyAsn 139
DB 730 ACAGCGTTGAGCTCTAGAGAAATTTGGAGGTGAACCGAGCACTATTTGGCCAGGAAC 789
QY 140 LeuGlyLeuLeuSerGlnLysValLysValLysValLysValLysValLysValLysValLys 159
DB 790 GCGAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 849
QY 160 ThrPheTrpLeuGlyThrLysLeuMetThrProGlyArgArgAlaIleTrpAlaIle 179
DB 850 TCATTCTACTGGGGAAACCAAGCTCATGACACCGAGAGCGCTTATGCTATCTGGGGGAT 909
QY 180 TyrValTrpCysArgArgTrpAspGlyLeuLeuValAspGlyProAsnAlaSerHisIleThr 199
DB 910 TATGTATGTGTAGAGGAGACAGATGAGCTTGTGATGGCCCTPACCGCTACACATTAAT 969
QY 200 ProGlnAlaLeuAspArgTrpGluTrpArgLeuGlnAspLysPheSerGlyArgProPhe 219
DB 970 CCAGCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1029
QY 220 AspMetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspLysGlnPro 239
DB 1030 GATATGCTTGAATGCTCTTATCTGATACATTCACCAAGTATCTGTGACATCCAGCA 1089
QY 240 PheArgAspMetIleGlyLysMetArgMetAspLeuTrpLysSerArgTrpLysThrPhe 259
DB 1090 TTTAGAGATATGATAGAGAAATGCGGATGATCTGAGAAATCGAGATACAAAGATTTC 1149
QY 260 AspGlyLeuTrpLeuTrpCysTrpTrpValAlaGlyThrValGlyLeuMetSerValPro 279
DB 1150 GATGAGCTGTATCTTACTGCTATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1209
QY 280 ValMetGlyIleAlaProGlySerLysValIleThrGlySerValLysValLysValLysVal 299
DB 1210 GTATATGCGATTCGACCTGATTAATTCAGCAACAGAAAGTGTATATGATGACAGCTT 1269
QY 300 AlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGlnAspAlaArg 319
DB 1270 TCTTTGGGATGCGGAGACGATGATTAATCTTAAGGAGCTTGAAGAGATGCAAGA 1329
QY 320 ArgGlyArgValLysLeuProGlnAspGlyLeuAlaGlnAlaGlyLeuSerAspGlyAsp 339
DB 1330 AAGAGGAGAGTGTACTTACTCAATGATTAAGCAAGCAAGCTTATCAGATGAGGAC 1389
QY 340 IlePheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArg 359
DB 1390 ATTTTTCGTAAGAAATTCACAGCAATGAGAGATTTTATGAGAGAGCAAAATCAAGG 1449
QY 360 AlaArgLysPhePheAspLeuSerGlyLysValLysValLysValLysValLysValLysVal 379
DB 1450 GCTAGAAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1509
QY 380 TrpPro---ValLeuThrAlaLeuLeuLeuLysValLysValLysValLysValLysVal 398
DB 1510 ATTGCTGTGTGGCCAGCGCTTGTGCTTTTATGAGAAATATTTGATGATGATGATGATGATG 1569

```

```

QY 399 AsnAspTrpAsnAsnPheThrArgArgAlaTrpValSerLysProLysLysLeuLeuThr 418
DB 1570 ATGACTACCAAAATTTACAAAAGGGCTTATGTTAAACAGGCAAGAGCTATTAGCT 1629
QY 419 LeuProIleAlaTrpAlaLysSerLeu 427
DB 1630 ATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1656
RESULT 14
AAV16949
ID AAV16949 standard; cDNA to mRNA; 1921 BP.
AC AAV16949;
XX
XX
XX 06-JUL-1998 (first entry)
XX
XX
XX Nucleic acid encoding phytoene synthase 2.
XX
XX Phytoene synthase; breeding; variable flower colour; ds.
XX
XX Gentiana lutea.
OS
XX Key Location/Qualifiers
FH CDS 412..1689
FT /*tag= a
XX
XX JPI0084966-A.
XX
XX 07-APR-1998.
XX
XX 17-SEP-1996; 96JP-0245107.
XX
XX 17-SEP-1996; 96JP-0245107.
XX
XX (IWAT-) IWATE KEN.
XX
XX PA
XX DR WPI; 1998-264853/24.
XX
XX P-PSDB; AAW46962.
XX
XX PT Phytoene synthase gene - useful for breeding plant of variable
XX flower colour
XX
XX PS Claim 2; Pages 7-9; 15pp; Japanese.
XX
XX CC The present sequence encodes phytoene synthase 2. It was isolated from
XX a cDNA library prepared from mRNA extracted from the petals of Gentiana
XX lutea. The nucleic acid sequence was amplified from the library using
XX CC PCR primers AAV16952-53. The phytoene synthase gene is useful for
XX breeding plants with variable flower colours.
XX
XX SQ Sequence 1921 BP; 633 A; 236 C; 438 G; 554 T; 0 other;
Alignment Scores:
Pred. No.: 3,42e-157 Length: 1921
Score: 1557.00 Matches: 305
Percent Similarity: 81.82% Conservative: 46
Best Local Similarity: 71.10% Mismatches: 68
Query Match: 68.59% Indels: 10
DB: 19 Gaps: 2
US-09-847-081B-2 (1-440) x AAV16949 (1-1921)
QY 1 MetSerMetSerValAlaLeuLeuTrpValValSerProThrSerGlyValSerAsnGly 20
DB 406 GTTAACATGCTATATTGTACGCTATGAGGTGTTTGGCCGAGTTCTGAAGTTTGTAGTGC 465
QY 21 ThrGlyLeuLeuAspSerValArgGlyLysAsnArgValPheValSerSerArgPheLeu 40
DB 466 AATGTTTCTTGAGCCCAATTCGAGAAAGT-----TACCAT 501
QY 41 AlaArgAspArgAsnLeuMetTrpAsnGlyArgGlyLysValLysValLysValLysValLysVal 60
DB 502 TTTTCGGATATAAAGTTTAAATGTCATGAGAGATTAGAGAAAGTACAGCAACCAAGCGCT 561

```

```

QY 61 AsnphgIysrLeuIleAlaAspProArgTyrSerCysLeu-----GlyGlySerArg 78
DB 562 AGATACGCTTATGGGCTTGAGATTGAGTTGATTCATTTGCTTGAGAGAGCTGGATTAG 621
QY 79 ThrGluYsGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGlu 98
DB 622 ACCCGGGAAGAAAGATTATCGGTATCCTCCAGATTAAATGATCAACCCCGGAGAGA 681
QY 99 MetThrValSerSerGluLeuValValTyrAspValValLeuValGlnAlaLeuVal 118
DB 682 ATGACGATGACATCAGACGAAAGGTTATGATGCTTTTAAACGACGCTTGATT 741
QY 119 LysArgGlnLeuArgSerThrAspAspLeuGluValLysProAspIleValValProGly 138
DB 742 AATGACAGCTTGAGCTTGAAGAAATTTGAGAGTGAACCGGACATTATTTGCAAGA 801
QY 139 AsnLeuGlyLeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAla 158
DB 802 AACCGGAACGTGTGATGAAGCTTATGTCGTCAGAAATGTCGTAATATGCTC 861
QY 159 LysThrPheTyrLeuGlyThrLeuMetThrProGluArgArgAlaIleTyrAla 178
DB 862 AAGTATTCTCTAGTGGGAAACGCTCATGACACCGAGAGCGCTTATGCTATCTGGCG 921
QY 179 IleTyrValITPcysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIle 198
DB 922 ATATATGATGTGTGATGAGAGACAGATGAGCTTGTGATGGCTTAAACCGCTCAACATA 981
QY 199 ThrProGlnAlaLeuAspArgTyrGluThrArgLeuGluAspIlePheSerGlyArgPro 218
DB 982 AATCCAAACCGCGTTAGATGATGAGTGGAGACCAAGATTAGAAAGTATTTCAAAGGCGCAACT 1041
QY 219 PheAspMetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGln 238
DB 1042 TTGTATATGCTTGAAGCTGCTTATCTGATACCATTAACCATATCTGTTGAGACATCAG 1101
QY 239 ProPheArgAspMetIleGluGlyMetArgMetAspLeuTyrLysSerArgTyrLysThr 258
DB 1102 CCATTATGAGATGATGATGAAGAGATGCGGATGATCTGAAAGAAATCCAGATACAGAGAT 1161
QY 259 PheAspGluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerVal 278
DB 1162 TTCGATGAGCTGTATCTTACTTACCTGATATATGCTGTCGATGCTGATGATGCTGTA 1221
QY 279 ProValMetGlyIleAlaProGluSerLysAlaThrThrGlnSerValTyrAsnAlaIle 298
DB 1222 CCAGTAAATGGGCAATGACCTGAATCTAAGGCAACAGAAAGTGTATATATACACT 1281
QY 299 LeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAla 318
DB 1282 TTATCTTTGGGATGCGCAACACAGCTGATCACTTTCAAGGAGCTTGGAGAAATGCA 1341
QY 319 ArgArgGlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGlu 338
DB 1342 AGAAGAGAGAGAGTGTACTTACCTCAAGTGAATTAGCAACAAGGCTTATCAATGAG 1401
QY 339 AspIlePheAlaGlyArgValThrAspLysTyrPheAsnPheMetLysLeuGlnIleGln 358
DB 1402 GACATTTTGTCTGGAAGAGTTACAGACAAATGAGAGATTTTATGAAGAGAAATCAAA 1461
QY 359 ArgAlaArgLysPhePheAspGlnSerGluYsGlyValIThGluLeuAspSerAlaSer 378
DB 1462 AGGCTTAAGAAATTTATATATATATGACAGAAAAAAGTCCCCCGAACTCAGCTCCGAGC 1521
QY 379 ArgTyrProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAla 398
DB 1522 AGATTGCGCTGTGGGACGCTTGTCTTTTATAGAAAAATATATGATGAGATAGTAAGCA 1581
QY 399 AsnAspTyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysIleLeuLeuThr 418
DB 1582 AATGACTACAAACAAATTTCAAAAAGGGGCTTATGTAAACMAAGCAAGAGCTATTAGCT 1641

```

```

QY 419 LeuProIleAlaTyrAlaLysSerLeu 427
DB 1642 ATGCTGTAGCATGTGCCAAGTCTCTC 1668

RESULT 15
AB212924
ID AB212924 standard; DNA, 1269 BP.
AC AB212924;
XX
XX 21-JAN-2003 (first entry)
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 729.
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX Arabidopsis thaliana.
XX
XX W0200216655-A2.
XX
XX 28-FEB-2002.
XX
XX 24-AUG-2001; 2001WO-US26685.
XX
XX 24-AUG-2000; 2000US-227866P.
XX 26-JAN-2001; 2001US-264647P.
XX 22-JUN-2001; 2001US-300111P.
XX
XX (SCRI ) SCRIPPS RES INST.
XX (SYGN ) STINGENTA PARTICIPATIONS AG.
XX
XX Harper JF, Krebs J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
XX
XX Identifying a stress condition to which a plant cell has been exposed
XX and producing plants with increased tolerance to these abiotic stresses
XX
XX Claim 144; SEQ ID NO 729; 577bp + Sequence Listing; English.
XX
XX The invention relates to identifying a stress condition to which a plant
XX cell has been exposed, comprising:
XX (a) contacting nucleic acid representative of expressed polynucleotides
XX in the plant cell with an array or probes representative of the plant
XX cell genome; and
XX (b) detecting a profile of expressed polynucleotides in the plant cell
XX characteristic of a stress response. The method is useful in the
XX production of transgenic plants, cells and seeds and in producing plants
XX with increased tolerance to abiotic stresses. The present sequence is that
XX of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
XX in methods of the invention.
XX Note: The sequence data for this patent is not represented in the printed
XX specification but is based on sequence information supplied to Derwent by
XX the European Patent Office.
XX
XX Sequence 1269 BP; 365 A; 230 C; 315 G; 359 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 7.71e-156 Length: 1269
XX Score: 1542.00 Matches: 318
XX Percent Similarity: 81.86% Conservative: 43
XX Best Local Similarity: 72.11% Mismatches: 51
XX Query Match: 67.93% Indels: 30
XX DB: 24 Gaps: 8
XX
XX US-09-847-081B-2 (1-440) x AB212924 (1-1269)
QY 1 MetSerMetSerValAlaLeuLeuTyrValVal-----SerProThrSerGluValSer 18
DB 1 ATGCTTCTTCTGTAGGACGCTTATGAGCTTCTTCTTAAATCCAGACCAATG 60
QY 19 AsnGlyThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArg 38

```

```

Db      61 AACCAATGGGCTTGTA-----AGGCTTCTAGATCTTCTAGCA 99
Qy      39 PheLeuAlaArgAspArgAsnLeuMetTrpAsnGlyArgIleValysGly-GlyArgG 58
Db      100 CTGTTCTCTCTCTTGTAG-----AATCAAGACTTAACAAGGTAAAGAAAG 147
Qy      58 nArgTrpAsnPhenGlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerAr 78
Db      148 CAGATACCAACTGGAGTCTTCT-----TTGTAAGAACCA 186
Qy      78 gThrGlyValysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGly 98
Db      187 AGTAGAAGAAAT-GGTGTT-----GTGCTTCAACTTAGTAGCAAGCTTCTCGACA 239
Qy      98 uMetThrValSerSerGlnValysValTyrAspValValLeuValAlaAlaLeuVal 118
Db      240 GATAGCTCTTCTCACTGAAGAGAGGTTTACAATGTGTGAACAAGCTGCTTGTGT 299
Qy      118 llyAsArgGlnLeuArgSerThrAsp--AspLeuGlnValIys-----ProAspIleVal 135
Db      300 GAACAACAAGCTAAGGCTCTTCTTATGACCTTGATGTAAGAAACCAAGATGTGT 359
Qy      135 lValProGlyAsnLeuGlyLeuLeuSerGlnAlaTyrAspArgCysGlyGlnValCysAl 155
Db      360 TCTTCTGGAGAGTTGAGTTGTGGTGAAGCTTATGATCGATGCGGTGAAGTTGCCG 419
Qy      155 aGluTyrAlaValThrPheTyrLeuGlyThrIlySleuMetThrProGlnArgArgAl 175
Db      420 TGAATATGCTAAGACCTTTATCTTGAACCTTGTATGACACCGAAAGGCGAAAGGC 479
Qy      175 aIleTrpAlaIleTyrValTrpCysArgArgThrAspGlnLeuValAspGlyProAsnAl 195
Db      480 GATTGGGCAATCTAGCTTTGGTGAAGAACTGATGAACCTTGTGATGGCCAAATGC 539
Qy      195 aSerHisIleThrProGlnAlaLeuAspArgTyrGlnThrArgLeuGlnAspIlePheSe 215
Db      540 TTCACATATTAACCCCATGGCTTTAGATAGTAGGAAGCAAGGTAGAAAGATCTTTCCG 599
Qy      215 rGlyArgProPheAspMetLeuAspAlaAlaLeuSerAspThrValSerArgPheProVa 235
Db      600 TGGTCGCTCTTGTGAATGCTGTGATGCTGCTCGCTGATACAGTTCAGATACCCGGT 659
Qy      235 lAspIleGlnProPheArgAspMetIleGlnGlyMetArgMetAspLeuTrpIysSerAr 255
Db      660 CGATATTCAGCACTTCGAGACATGATCGAAGAAATGACATGACTTGAAGAAATCGAG 719
Qy      255 gTyrIlyThrPheAspGlnLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLe 275
Db      720 ATACCAAGACTTCGATGATCTATACCTTACTGCTACTAGCTGCGTGAACGCTCGATT 779
Qy      275 uMetSerValProValMetGlyIleAlaProGlnSerIlyAlaThrThrGlnSerValTy 295
Db      780 GATGAGCGTTCGCGTATGGAATCGATCTTAAGTGAAGCAACCAACGAAAGTGTTA 839
Qy      295 rAsnAlaAlaLeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValAl 315
Db      840 CAACGCTGCTTGCCCTTGCTGTATGACCAATCAGCTTACTACATACCTCGAAGCGTAG 899
Qy      315 yGlnAspAlaArgArgGlyArgValTyrLeuProGlnAspGlnLeuAlaGlnAlaGlyLe 335
Db      900 CGAAGATGCGAAGAGGAAAGGAGGTTATCTGCTCGAGATGAATGGCTCAGGCTGCT 959
Qy      335 uSerAspGlnAspIlePheAlaGlyArgValThrAspIlyGTrpArgAsnPhenMetIlyS 355
Db      960 TTCAGATGAAGCAATTTGCGCGAAAGTAACTGATTAATGAGAAACCTTCAATGAATAA 1019
Qy      355 sGlnIleGlnArgAlaArgIlyPhePheAspGlnSerGlnValysGlyValThrGlnLeuAs 375
Db      1020 GAGCTTAAACGACAAAGATGTTCTTCAGACGAGCTGAAGAAAGCGCTCACGAGCTCAG 1079
Qy      375 pSerAlaSerArgTyrProValLeuThrAlaLeuLeuLeuTyrArgIlySileuAspGln 395

```

```

Db      1080 TCCCGTAGACAGATGCGCTGTATGGCTTCAATTGCTATTGTACAGAGAAATCTGACGA 1139
Qy      395 uIleGlnAlaAsnAspTyrAsnAsnPhenThrArgArgAlaTyrValSerIysProIlyS 415
Db      1140 GATTGAAGCAATGATTAACAACAAATTTTACTAAGAGAGCTTATGTGGGAAAGTCAAGAA 1199
Qy      415 sLeuLeuThrLeuProIleAlaTyrAlaIysSerLeuValProProAsnArgThrSerSe 435
Db      1200 AATTGAGCTTGGCATTTGGCTTATGCTTAATCACTACTA-----AAGACTTCAAG 1250
Qy      435 T 435
Db      1251 T 1251

```

Search completed: January 16, 2004, 06:36:18  
 Job time : 389 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 16, 2004, 06:07:58 ; Search time 83 Seconds  
(without alignments)  
2339.862 Million cell updates/sec

Title: US-09-847-081b-2  
Perfect score: 2270  
Sequence: 1 MSMSVALLMVSPSTSEVNSG.....IAYANSLVPNRTSSPLAKT 440

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 569978 seqs, 220691566 residues  
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cg2\_1/USPTO\_SPOOL/US09847081/runat.15012004.154337.1879/app.query.fasta.1.583  
-DB=Issued\_Patents\_NA -QFMT=fasted -SUFFIX=rn1 -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORExt -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09847081 @CGN 1.1.56 @runat.15012004.154337.1879 -NCPU=6 -ICPU=3  
-NO MAP -LARGEBUTTER -NEG\_SCORES=0 -WAIT -DSFBLOCK=100 -LONELOG  
-BEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cg2\_6/prodata/2/ina/5A\_COMB.seq:\*  
2: /cg2\_6/prodata/2/ina/5B\_COMB.seq:\*  
3: /cg2\_6/prodata/2/ina/6A\_COMB.seq:\*  
4: /cg2\_6/prodata/2/ina/6B\_COMB.seq:\*  
5: /cg2\_6/prodata/2/ina/6CTUS\_COMB.seq:\*  
6: /cg2\_6/prodata/2/ina/backfillseq1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1826	80.4	1826	1	US-08-579-667-5 Sequence 5, Appl1
2	1799	79.3	1814	1	US-08-579-667-7 Sequence 1, Appl1
3	1774	78.1	1795	1	US-08-579-667-1 Sequence 1, Appl1
4	1762	77.6	1316	1	US-08-579-667-3 Sequence 3, Appl1
5	1750	77.1	1239	3	US-09-180-342-2 Sequence 2, Appl1
6	1744	76.8	1239	3	US-09-180-342-1 Sequence 1, Appl1
7	1732	76.3	1646	1	US-07-995-950-2 Sequence 2, Appl1
8	1732	76.3	1646	1	US-08-300-582-2 Sequence 2, Appl1
9	442	19.5	749	1	US-08-579-667-9 Sequence 9, Appl1
10	364	16.0	1198	1	US-08-995-726-5 Sequence 5, Appl1
11	364	16.0	1198	1	US-08-096-043-5 Sequence 5, Appl1
12	364	16.0	1198	1	US-08-093-577-5 Sequence 5, Appl1

13	364	16.0	1198	1	US-08-096-623A-5 Sequence 5, Appl1
14	361.5	15.9	1083	1	US-08-331-004A-1 Sequence 1, Appl1
15	361.5	15.9	1083	5	PCT-US95-13937A-1 Sequence 1, Appl1
16	350.5	15.4	1232	1	US-08-908-758-1 Sequence 1, Appl1
17	349.5	15.4	6818	1	US-07-783-705A-13 Sequence 13, Appl1
18	338.5	14.9	891	1	US-07-783-705A-11 Sequence 11, Appl1
19	334	14.7	4403765	3	US-09-103-840A-1 Sequence 2, Appl1
20	334	14.7	4411529	3	US-09-103-840A-1 Sequence 1, Appl1
21	326	14.4	8625	3	US-08-980-833-1 Sequence 27, Appl1
22	326	14.4	11233	3	US-08-980-833-27 Sequence 27, Appl1
23	324	14.3	908	3	US-08-660-645A-4 Sequence 4, Appl1
24	324	14.3	908	3	US-09-298-718-4 Sequence 4, Appl1
25	324	14.3	908	3	US-09-546-963-4 Sequence 4, Appl1
26	293	12.9	536165	4	US-09-214-808-1 Sequence 1, Appl1
27	176.5	7.8	947	1	US-08-096-623A-16 Sequence 16, Appl1
28	176.5	7.8	947	1	US-08-096-623A-17 Sequence 17, Appl1
29	174	7.7	2470	4	US-09-091-725-18 Sequence 18, Appl1
30	174	7.7	2546	4	US-09-091-725-12 Sequence 12, Appl1
31	170	7.5	405	3	US-09-060-756-345 Sequence 345, App
32	170	7.5	405	4	US-09-670-314-345 Sequence 345, App
33	159	7.0	3550	4	US-09-091-725-22 Sequence 22, Appl1
34	136.5	6.0	2054	1	US-08-351-981-1 Sequence 1, Appl1
35	121	5.3	1349	1	US-08-351-981-5 Sequence 5, Appl1
36	120	5.3	1326	4	US-09-625-188-3 Sequence 3, Appl1
37	116	5.1	1642	1	US-08-310-693-1 Sequence 1, Appl1
38	116	5.1	1642	5	PCT-US95-11280-1 Sequence 10, Appl1
39	112.5	5.0	114	4	US-08-260-546-10 Sequence 12, Appl1
40	112.5	5.0	114	4	US-09-436-068A-12 Sequence 12, Appl1
41	112.5	5.0	114	4	US-09-265-576-12 Sequence 427, App
42	111.5	4.9	346	3	US-09-060-756-427 Sequence 427, App
43	111.5	4.9	346	4	US-09-670-314-427 Sequence 133, App
44	108	4.8	39195	4	US-08-311-731A-133 Sequence 545, App
45	107.5	4.7	425	3	US-09-060-756-545

## ALIGNMENTS

RESULT 1  
US-08-579-667-5  
Sequence 5, Application US/08579667  
Patent No. 5705624  
GENERAL INFORMATION:  
APPLICANT: Filizmaurice, Wayne P.  
APPLICANT: Hellmann, Gary M.  
APPLICANT: Grill, Laurence K.  
APPLICANT: Kungel, Monto H.  
TITLE OF INVENTION: DNA SEQUENCES ENCODING ENZYMES USEFUL IN PHYTOENE BIOSYNTHESIS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Virginia C. Bennett  
STREET: 1211 East Morehead Street, PO Drawer 34009  
CITY: Charlotte  
STATE: No. 5705624th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/579,667  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Bennett, Virginia C.  
REGISTRATION NUMBER: 37,092  
REFERENCE/DOCKET NUMBER: 627-196  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-420-2200  
TELEFAX: 919-881-3175



```

/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1814 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 363..1592
US-08-579-667-7

Alignment Scores:
Pred. No.: 4,71e-201 Length: 1814
Score: 1799.00 Matches: 356
Percent Similarity: 87.93% Conservative: 30
Best Local Similarity: 81.09% Mismatches: 24
Query Match: 79.25% Indels: 29
DB: 1 Gaps: 4

US-09-847-081b-2 (1-440) x US-08-579-667-7 (1-1814)

QY 3 MetSerValAlaLeuLeuTyrValValSerProThrSerGluValSerAsnGlyThrGly 22
DB 363 ATGTCTGTTGCTTGTATGAGGTGTTCACCT--TGTGAAGCTCCAAATGGACAGCA 419
QY 23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42
DB 420 TTCTTGATTCAGTCCGGAGGGAACCGGGTTTGTGATTCGTC-----AGG 467
QY 43 AspArgAsnLeuMetTyrAsnGlyArgGlyLeuGlyGlyArgGlnArgTyrAsn 62
DB 468 CATAGAAATTATGTGTGCAATGAGAAACAGAGAGGTGTGAACAAAGTGGAATTTT 527
QY 63 GlySerLeuLeuAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluGly 82
DB 528 GGT----- 530
QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMet---ThrVal 101
DB 531 -----TCTGTAAGGCTCTGTAAGGTGCTACACCGCGGAGAAATGGCGAGAG 581
QY 102 SerSerGluValLeuValTyrAspValValLeuValGlnAlaLeuValIysArgGln 121
DB 582 ACATCAGAACAGATGTTATGATGTGCTTTAAACAGACGCTTATGTAAGAGCGAG 641
QY 122 LeuArgSerThrAspAspLeuGluValIysProAspIleValValProGlyAsnLeuGly 141
DB 642 TTGAATCTGCTGATGATTTAGAGTGAAGCCGAGATCCCTCCCGGAAATTTGAGC 701
QY 142 LeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaIysThrPhe 161
DB 702 TTGTTGAGGAGACATGATGATGAGTGTAGTGTGAGATGTCAGAGATGCCAAGACATT 761
QY 162 TyrLeuGlyThrIysLeuMetThrProGluArgArgAlaIleTyrAlaIleTyrVal 181
DB 762 TACTTAGAGAACCATGTAATGATCCAGAGAGAAAGGCTATTGGGCATATATGTG 821
QY 182 TyrCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGln 201
DB 822 TGGTCAGAGAGAACATGATGATCTGTTGATGCGCCCAAGCATCATGTTACACCCCA 881
QY 202 AlaLeuAspArgThrGluThrArgLeuGluIlePheSerGlyArgProPheAspMet 221
DB 882 GCTTAGAGATGGTGGAGAACCGGCTTGAAGATGTTTTCAGGGGGACATTGTATG 941
QY 222 LeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArg 241
DB 942 CTCGATGCTGCTTGTGCCATGCTGTTCCAGCTTCCAGCTTCAATATGACCGCTTCA 1001
QY 242 AspMetIleGluGlyMetArgMetAspLeuTyrIysSerArgTyrIysThrPheAspGlu 261
DB 1002 GATATGATTTGAAGAAATGCGTATGACTTGAGGAAGTCAAGATATATGAACCTTTGATGAG 1061

```

```

QY 262 LeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMet 281
DB 1062 CTTTACCTCTATGTTATTAACCTTCTGCTACGCTGGGCTTGATGATGATTCATTTATG 1121
QY 282 GlyIleAlaProGluSerIysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeu 301
DB 1122 GGTATTTGACCTGATTCAAAGGCAACACAGAGACGTATATATGACGCTTGCTTTA 1181
QY 302 GlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGly 321
DB 1182 GGAATCGGAATCAACTACAGCAACTACTCAGGAGTGTGAGAAAGATGCCAGAAAGAGA 1241
QY 322 ArgValTyrLeuProGluAspGluLeuValGlnAlaGlyLeuSerAspGluAspIlePhe 341
DB 1242 AGAGCTCTATTAACCTCAAGATGATTAATGACACAGGAGGCTCTTGCAGATGACATATTT 1301
QY 342 AlaGlyArgValThrAspIysTyrAsnProMetIysGlnIleGlnArgAlaArg 361
DB 1302 GCTGGAAGATGACATGATAGTGTGAGAGAGCTTTATGAGAAGCAATCCAGAGGCAAGA 1361
QY 362 LysPhePheAspGluSerGluGlyValThrGluLeuAspSerAlaSerArgTyrPro 381
DB 1362 AAGTCTTCGATGAGGACAGAGAGGAGGAGTTCACAACTGAGCTCAGCTAGACAGATGGCT 1421
QY 382 ValLeuThrAlaLeuLeuLeuTyrArgIysIleLeuAspGluIleGluAlaAsnAspTyr 401
DB 1422 GTATGGCATCTTCTGCTGTTGACCGCAATATCTGAGACGAGATTGAAGCCATATACATC 1481
QY 402 AsnAsnProThrArgArgAlaTyrValSerIysProIysIysLeuLeuThrLeuProIle 421
DB 1482 AACAACTTCACAAAGAGCTTATGTGAGCAAAACCAAGAAAGCTAATTCCTTACCTAAT 1541
QY 422 AlaTyrAlaIysSerLeuValProProAsnArgThr-SerSerProLeuAlaIys 439
DB 1542 GCTATGCAAAATCTTGTGCCCCCTTCAAGAACTTGTCACTTACCTTGAAG 1596

RESULT 3
US-08-579-667-1
/ Sequence 1, Application US/08579667
/ Patent No. 5705624
/ GENERAL INFORMATION:
/ APPLICANT: Filizmaurice, Wayne P.
/ APPLICANT: Hellmann, Gary M.
/ APPLICANT: Grill, Laurence K.
/ APPLICANT: Kumagai, Monto H.
/ APPLICANT: Della-Cioppa, Guy R.
/ TITLE OF INVENTION: DNA SEQUENCES ENCODING ENZYMES USEFUL IN
/ NUMBER OF SEQUENCES: 19
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Virginia C. Bennett
/ STREET: 1211 East Morehead Street, PO Drawer 34009
/ CITY: Charlotte
/ STATE: No. 5705624th Carolina
/ COUNTRY: USA
/ ZIP: 28234
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/579,667
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bennett, Virginia C.
/ REGISTRATION NUMBER: 37,092
/ REFERENCE/DOCKET NUMBER: 627-196
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 919-420-2200
/ TELEFAX: 919-881-3175

```





```

REFERENCE/DOCKET NUMBER: 627-196
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1316 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1239
US-08-579-667-3

Alignment Scores:
Pred. No.: 6,186-197 Length: 1316
Score: 1762.00 Matches: 347
Percent Similarity: 87.04% Conservative: 29
Best Local Similarity: 80.32% Mismatches: 28
Query Match: 77.62% Indels: 28
DB: 1 Gaps: 4

US-09-847-081b-2 (1-440) x US-08-579-667-3 (1-1316)
QY 3 MetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThrGly 22
Db 1 ATGCTGTTGCTTGTATGAGTGTGTTTCACT---TGGAGGCTTCAATGGACAGGA 57
QY 23 LeuLeuAspSerValArgGluGluValAsnArgValPheValSerSerArgPheLeuAlaArg 42
Db 58 TTCTTGATTCATCCGGAGGGAACCGGGTTTGGATGTCG-----AGG 105
QY 43 AspaGanLeuMetTrpAsnGlyArgGlyLeuValGlyValArgGlnArgTrpAsn 62
Db 106 CATAGAAATTGATGTCATAGAGAAACAGAGAGGTGCGAACAAATGTGGAATTT 165
QY 63 GlySerLeuLeuAlaAspProArgTrpSerCysLeuGlyGlySerArgThrGluValGly 82
Db 166 GGT----- 168
QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMet---ThrVal 101
Db 169 -----TCTGTAAGCTCTGTATGAGTGGCTACACCGCGGGAATGGCAGATG 219
QY 102 SerSerGluValValValValValValValValValValValValValValValValVal 121
Db 220 ACATCAGAACAGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 279
QY 122 LeuArgSerThrAspAspLeuGluValValValValValValValValValValValValVal 141
Db 280 TTGACATCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 339
QY 142 LeuLeuSerGluValValValValValValValValValValValValValValValValVal 161
Db 340 TTGTTGATGGAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 399
QY 162 TyrLeuGlyThrValLeuMetThrProGluArgArgValAlaIleTyrAlaIleTyrVal 181
Db 400 TACTTAGAGAACTATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 459
QY 182 TrpCysArgArgTrpAspGluLeuValAlaAspGlyProAlaAlaSerHisIleThrProGln 201
Db 460 TGGTGACAGAGAACAGCAAGAACTTGTGATGCGCCGAAAGCAATCAATTAATCTCACAA 519
QY 202 AlaLeuAspArgTrpGluThrArgLeuGluValAlaIlePheSerGlyArgProPheAspMet 221
Db 520 GCTTTAGATAGGTGGGAAGACCGGCTGGAGAGATGTTTCAAGTGGCGGCATTTACATG 579
QY 222 LeuAspAlaAlaLeuSerAspThrValValSerArgPheProValAspIleGlnProPheArg 241
Db 580 CTCGATGCTGCTTTGCTCGATGATGCTTTCCAAAGTTTCCAGTTGATATTCAGCCGTTCA 639

```

```

QY 242 AspMetIleGluGlyMetArgMetAspLeuTrpValSerArgTrpValSerThrPheAspGlu 261
Db 640 GATATGATCGAAGAAATCCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 699
QY 262 LeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMet 281
Db 700 CTATACCTATATGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 759
QY 282 GlyIleAlaProGluSerIleValSerThrGluSerValTyrAsnAlaAlaLeuAlaLeu 301
Db 760 GGCATGCACTGAAATCAAGGCAACACAGAGATGTTATATATGACAGCTTGGCTTTG 819
QY 302 GlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGly 321
Db 820 GGTATCCGGAATCAACTAACCAACTTCTCAGAGATGTCGAGAAATGCCAGAAAGAGA 879
QY 322 ArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePhe 341
Db 880 AGAGCTCACTTAACCTCAAGATGAATTAACACAGGAGGCTCTCCGACGATGACATATT 939
QY 342 AlaGlyArgValThrAspValTrpArgAsnPheMetIleValGlnIleGlnArgAlaArg 361
Db 940 ACTGGAAGAGTACGATTAATGAGAGAGCTTTATGAAAGACATTCAGAGGCAAGA 999
QY 362 LysPhePheAspGluSerGluValGlyValThrGluLeuAspSerAlaSerArgTrpPro 381
Db 1000 AAGTTCTTCAATGAGGCAAGAGAAAGATTACACACTGAGCTCAGCTACGATGAGCT 1059
QY 382 ValLeuThrAlaLeuLeuLeuTyrArgValIleLeuAspGluIleGluAlaAsnAspTyr 401
Db 1060 GTATGGGATCTTGTCTGTGTGACCGCAATATCTCGAGATGAGAGCAATGATAC 1119
QY 402 AsnAsnPheThrArgArgAlaTyrValSerIleProValValValValValValValVal 421
Db 1120 AACCACTTCAACAAAGAGGCTTATGTCGCAAAATCAAGAAAGCTAATTCCTTACCTATT 1179
QY 422 AlaTyrAlaValSerLeuValProProAsnArgThr 433
Db 1180 GCTTATGCAAAATCTTGTGTGCCCCCTTACAGAACT 1215

```

```

RESULT 5
US-09-180-342-2
; Sequence 2, Application US/09180342A
; Patent No. 6239331
GENERAL INFORMATION:
; APPLICANT: Drake, Caroline R.
; APPLICANT: Bird, Colin R.
; APPLICANT: Schuch, Wolfgang W.
; TITLE OF INVENTION: Enhancement of Gene Expression
; FILE REFERENCE: SE850156
; CURRENT APPLICATION NUMBER: US/09/180,342A
; CURRENT FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: PCT/GB97/01414
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: UK 9611981.3
; EARLIER FILING DATE: 1996-06-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
US-09-180-342-2

```

```

Alignment Scores:
Pred. No.: 1,446-195 Length: 1239
Score: 1750.00 Matches: 344
Percent Similarity: 86.84% Conservative: 32
Best Local Similarity: 79.45% Mismatches: 33
Query Match: 77.09% Indels: 24
DB: 3 Gaps: 3

```

US-09-847-081B-2 (1-440) x US-09-180-342-2 (1-1239)

QY 3 MetSerValAlaLeuLeuTrpValValSerProThrsSerGluValSerAsnGlyThrGly 22  
 DB 1 ATGCTGTGCTGTTATGGGTTGTTCTCT--TGTGACGTCCTAAATGGGCAAGT 57  
 QY 23 LeuLeuAspSerValArgGluGluYAsnArgValPheValSerSerArgPheLeuAlaArg 42  
 DB 58 TTCAATGCAATGATCGGAGGAGAAACGCTTTTGTGATTCATCG-----AGG 105  
 QY 43 AspArgAsnLeuMetTrpAsnGlyArgGlyLeuSlySlyGlyGlyArgGlnArgTrpAsnPhe 62  
 DB 106 CATGGCAATTTGCTCCAAATAGAGAAATCAATAGAGCT-----144  
 QY 63 GlySerLeuLeuAlaAspProArgTrpSerCysLeuGlyGlySerArgThrGluSlyGly 82  
 DB 145 -----GGTGGAAAGCAAACTAATATATGA 168  
 QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThrValSer 102  
 DB 169 CGGAATTTTCTGATCGCTGCTATTTGGTACTCATCTGAGAGAACGAGCATGACA 228  
 QY 103 SerGluSlySlyValTyrAspValValLeuSlyGlnAlaAlaLeuValLysArgGlnLeu 122  
 DB 229 TCGAACAAGATGCTATGATGATGCTTTTGAGGACAGCGCTTGTAAGAGGCAACTG 288  
 QY 123 ArgSerThrAspAspLeuGluValLysProAspIleValProGluYAsnLeuGlyLeu 142  
 DB 289 AGATCTACCAATGAGTTAGAGTGAAGCCGAGATACCTATTCGAGGAAATTTGGCTTG 348  
 QY 143 LeuSerGluAlaTyrAspArgCysGlyGlyValCysAlaGluTyrAlaLysThrPheTyr 162  
 DB 349 TTGAGTGAAGCATATGATAGGTGCTGTAAGATGTGACAGTATGCAAAAGCGTTTAC 408  
 QY 163 LeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrpAlaIleTyrValTrp 182  
 DB 409 TTAGGAACTATGCTATGATGCTCCAGAGAAAGGCGTATCTGGCAATATATGTATG 468  
 QY 183 CysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGlnAla 202  
 DB 469 TGCAGAAACAGATGAACTGTTGATGCCCCAAACGATCATATATATATATATATATAT 528  
 QY 203 LeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeu 222  
 DB 529 TTAGTATAGTGGGAAATAGGCTAGAGATGTTTCANTGGCGCCATTTGACATGCTC 588  
 QY 223 AspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArgAsp 242  
 DB 589 GATGCTGCTTTGCTCCGATACAGTTTCTAACTTTCAGTTGATATTCACCCATTCAGAGAT 648  
 QY 243 MetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysThrPheAspGluLeu 262  
 DB 649 ATGATTAAGAGAAATGCTATGACTTGAAGAAATGAGAAATCAAAAACCTTCGACCAACTA 708  
 QY 263 TyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGly 282  
 DB 709 TACCTTATTTGTTATATGCTGCTGTAACGCTGGGTTGATGATGATGCTTCCATATATG 768  
 QY 283 IleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGly 302  
 DB 769 ATCCGCCCTGATCAAAAGCAACACAGAGCGTATATATGCTGCTTGGCTCTGGGG 828  
 QY 303 LeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGlyArg 322  
 DB 829 ATCCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 888  
 QY 323 ValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAla 342  
 DB 889 GTCTACTGCTCAAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 948  
 QY 343 GlyArgValThrAspLysTrpArgAsnPheMetLysGlyGlnIleGlnArgAlaArgLys 362  
 DB 949 GGAAGGTGACCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1008

QY 363 PhePheAspGluSerGluSlyGlyValThrGluLeuAspSerAlaSerArgTrpProVal 382  
 DB 1009 TTCTTATATAGAGAGAGAAAGGCTGACAGATTAAGCTCAGTATATATATATATATAT 1068  
 QY 383 LeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAspTyrAsn 402  
 DB 1069 TGGGATCTTTGGCTTGTATACCGCAAAATATAGATAGATTAAGCAATGATCAAC 1128  
 QY 403 AsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIleAla 422  
 DB 1129 AACTTCAAAAGAGATATATGCAAAATCAAAAGTTGATGCAATTAATTAATTAAT 1188  
 QY 423 TyrAlaLysSerLeuValProProAsnArgThrSerSer 435  
 DB 1189 TATGCAAAATCTCTTGCTGCTCTTCAAAAACTGCTCT 1227

RESULT 6  
 US-09-180-342-1  
 ; Sequence 1, Application US/09180342A  
 ; Patent No. 6239331  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Drake, Caroline R.  
 ; APPLICANT: Bird, Colin R.  
 ; APPLICANT: Schuch, Wolfgang W.  
 ; TITLE OF INVENTION: Enhancement of Gene Expression  
 ; FILE REFERENCE: SEE50156  
 ; CURRENT APPLICATION NUMBER: US/09/180,342A  
 ; CURRENT FILING DATE: 1998-11-05  
 ; EARLIER APPLICATION NUMBER: PCT/GB97/01414  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: UK 9611981.3  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 1239  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA  
 ; US-09-180-342-1

Alignment Scores:  
 Pred. No.: 7,27e-195 Length: 1239  
 Score: 1744.00 Matches: 343  
 Percent Similarity: 86.61% Conservative: 32  
 Best Local Similarity: 79.21% Mismatches: 34  
 Query Match: 76.83% Indels: 24  
 DB: 3 Gaps: 3

US-09-847-081B-2 (1-440) x US-09-180-342-1 (1-1239)

QY 3 MetSerValAlaLeuLeuTrpValValSerProThrsSerGluValSerAsnGlyThrGly 22  
 DB 1 ATGAGCTGGGCACTTCTTGGTGGTGAAGCCCA---TGCATGTGATTAACGGCAGCTTCA 57  
 QY 23 LeuLeuAspSerValArgGluGluYAsnArgValPheValSerSerArgPheLeuAlaArg 42  
 DB 58 TTTATGAGAGAGTGAAGAGGTAATATGATTTTCAAGCTTCT-----CGT 105  
 QY 43 AspArgAsnLeuMetTrpAsnGlyArgGlyLeuSlySlyGlyGlyArgGlnArgTrpAsnPhe 62  
 DB 106 CACCGTAACTTGTATGATTAACGAACGTAATTAACAGGGCA-----CGT 144  
 QY 63 GlySerLeuLeuAlaAspProArgTrpSerCysLeuGlyGlySerArgThrGluSlyGly 82  
 DB 145 -----GAGGTAAACAGACAAACAAACAGCGT 168  
 QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThrValSer 102  
 DB 169 AGAAAGTTCTCAAGTTAGATGACGAATCTTGGCAACACCTGAGCGGTGAGAGAACTATGACT 228

OY	103	SerGluValValValTyrAspValValLeuValSerGluAlaValLeuValLysArgGluLeu	122
Db	229	AGCGAGCAATGGTGTACACGCTCGTACCTTCTTCAGAGCTGCACCTGTTAAACGTACGTTA	288
OY	123	ArgSerThrAspAspLeuGluValLysProAspIleValValProGluYasnLeuGluLeu	142
Db	289	CGTAGTACTTAAACGAACTTGAGGTTAACTGACATTCGAACTCTGGAAACCTTGGACTT	348
OY	143	LeuSerGluAlaTyrAspArgCysGluValCysAlaGluTyrAlaLysThrPheTyr	162
Db	349	CTTCTGTAGGCTTACGACAGATGGCGAGAGGTTTGGCGAAATACGTTAAACCTTCAAT	408
OY	163	LeuGluTyrLysLeuMetThrProGluArgArgAlaIleThrAlaIleTyrValTyr	182
Db	409	TTGGGTACCATGTGTGTACACAGAAAGGCGCTGGCAATATGGGCTATTTACGTTGG	468
OY	183	CysArgArgThrAspGluLeuValAspGluProAsnAlaSerHisIleThrProGluAla	202
Db	469	TGTAGCGCGTACGACGAGTTAGTACGCGACCTATGCTAGTTACTTAACACCCCGCTGT	528
OY	203	LeuAspArgTyrGluThrArgLeuGluAspIlePheSerGluYArgProPheAspMetLeu	222
Db	529	CTTGCACAGTGGGAAACCGTTTGGAGAGCGGTTTAAACGGACAGCTTTCATATGTTG	588
OY	223	AspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGluProPheArgAsp	242
Db	569	GACGGAGCACTTAAGTACACTGTGAGCAATTTCCCTGGAGCATCCMACCTTTTGGGAC	648
OY	243	MetIleGluGluMetArgMetAspLeuTyrLysSerArgTyrLysThrPheAspGluLeu	262
Db	649	ATGATCGAGGCGATAGAAATGATCTTCGTAAGTCTCGTTATAGAAATTTGATGATTTG	708
OY	263	TyrLeuTyrCysTyrTyrValAlaGluThrValGluLeuMetSerValProValMetGly	282
Db	709	TATTTGTACTGCTACTACGTCGACAGAAACGCGGGCGTTATGTCAGTGCCTTACATGGGA	768
OY	283	IleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGly	302
Db	769	ATTGACACAGAAAGTAAAGCTACTGATGATGTTTACACCGCAGCACTTACCATTTAGGT	828
OY	303	LeuAlaAsnGluLeuThrAsnIleLeuArgAspValGluAspAlaArgArgGlyArg	322
Db	829	ATAGCTAACCACTTACAAATATCTTGAAGGACGCGGGTGAAGACGACGTAAGGGCTGCT	888
OY	323	ValTyrLeuProGluAspGluLeuAlaGluAlaGluLeuSerAspGluAspIlePheAla	342
Db	889	GGTATCTCTCCACAGACGAGCTCGCTCAAGCTGGATGAGTGAAGAGAGACATTTTCGA	948
OY	343	GluArgValThrAspLysTyrPheAsnPheMetLysGluIleGluArgAlaArgLys	362
Db	949	GGTCCGTGTACAGACAAAGTGAAGATTTTCAATGAAAGCAAGATTCACCGTCTCGTAA	1008
OY	363	PhePheAspGluSerGluLysGluValThrGluLeuAspSerAlaSerArgTyrProVal	382
Db	1009	TTTTTCGACGAAGCTGAAGAGGAGCTTACTGAGCTTCTTACGTCATCAAGGTTTCCAGTT	1068
OY	383	LeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAspTyrAsn	402
Db	1069	TGGCGCACCGCTTGCTCTATAGAAAGATTTTGGACGAATCGAGGCTTACATATATAT	1122
OY	403	AsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIleAla	422
Db	1129	AATTTTACTTAAACGCGCTTACCTTCTTAAAGACAAAAAACCTTATGCTCTTCCAAATCGCT	1188
OY	423	TyrAlaLysSerLeuValProProAsnArgThrSerSer	435
Db	1189	TACGCTAAGAGCTTGCTTCCACCACTAAGACAGCTTAC	1227

RESULT 7  
US-07-995-950-2  
; Sequence 2, Application us/07995950  
; Patent No. 5304478  
; GENERAL INFORMATION:

[illegible]

```

Db 423 ATGACATCGAACAACAGTGTATGATGCTTTGAGCGACGACCTTGCTGAAGAG 482
Qy 121 GlnleuAysSerThrAspAspLeuGluValLysProAspIleValProGlySerLeu 140
Db 483 CAATGAGATCTTACCAAGAGTTAAGAGGAGCGGAGATACCTATTCCGGGGAAATTG 542
Qy 141 GlyLeuLeuSerGluValATyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThr 160
Db 543 GCGCTTGAGTGAAGCATATGATAGTGTTGAGTAAGTATGTCAGAGATGCAAGAGCG 602
Qy 161 PheTyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTyrAlaIleTyr 180
Db 603 TTTAACTTAGGAACTATGCTATGATGCTCCGAGAGAGAGAGGCTATCTGGGCAATATAT 662
Qy 181 ValTyrCysArgArgThrAspGluLeuValAspGlyProAspAlaSerHisIleThrPro 200
Db 663 GTATGTCAGAGAGAGAGAGATGATGCTGTTGATGCGCCAAAGCATATATATATATAT 722
Qy 201 GluAlaLeuAspArgTyrGluThrArgLeuGluAspIlePheSerGlyArgProPheAsp 220
Db 723 GCAGCTTAGATAGTGGGAGAAATAGCTAGAGAGATGTTTCAATGGCGCGCATTTGAC 782
Qy 221 MetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPhe 240
Db 783 ATGCTCGATGGGCTTGTGTCGATACAGTTCTAATCTTCCAGTATATTCAGCCATTC 842
Qy 241 ArgAspMetIleGluGlyMetArgMetAspLeuTyrPylSerArgTyrLysThrPheAsp 260
Db 843 AGAGATATGATTTGAAGAGATGGATGAGCTTGAGAGAAATCAGATATCAAAAATCTTCAC 902
Qy 261 GluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProVal 280
Db 903 GACATATACCTTATATGTTATATTTGCTGCTGACGTTGGTTGATGAGTCTTCCATT 962
Qy 281 MetGlyIleAlaProGlySerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAla 300
Db 963 ATGGGTATCGCCCTGGAATCAAAAGGCAACAACAGAGAGATATATATGCTTGGCT 1022
Qy 301 LeuGlyLeuAlaAsnGluLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArg 320
Db 1023 CTGGGATCGCAAAATCAATTAACATACTCAAGATGTTGAGAGAGATGCCAGAGA 1082
Qy 321 GlyArgValTyrLeuProGlnAspGluLeuAlaGlyLeuSerAspGluAspIle 340
Db 1083 GGAAGAGCTTACTTCTCAGATGATATTAAGCAGCAGGCTCTATCCGATGAGATATA 1142
Qy 341 PheAlaGlyArgValThrAspLysTyrArgAsnPheMetLysLeuGlnIleGlnArgAla 360
Db 1143 TTTGCTGGAAGGATGACGATTAATGAGAAATCTTATGAGAGAAACAATACATAGGCA 1202
Qy 361 ArgLysPhePheAspGluSerGlyLysGlyValThrGluLeuAspSerLysSerArgTyr 380
Db 1203 AGAAAGTCTTGTGAAGAGCAGAGAAAGCCGAGCAAGATTGACCTAGATGATGATGAT 1262
Qy 381 ProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAsp 400
Db 1263 CCTGATAGGCGATCTTGTGCTTGTGACCGCAAAATACATGATGATGATGATGATGAT 1322
Qy 401 TyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuPyr 420
Db 1323 TACAACAACCTTCAACAAGAGAGCATATGATGAGCAAAATCAAGCAAGTATGATGATGAT 1382
Qy 420 CilleAlaTyrAlaLysSerLeuValProProAsnArgThrSerSerProLeuAlaLys 439
Db 1383 TATGCAATATGCAAAATCTTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1434

```

RESULT 8  
 US-08-300-582-2  
 ; Sequence 2, Application US/08300582  
 ; Patent No. 5750865  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BIRD, COLIN R.

```

; APPLICANT: GRIERSON, DONALD
; APPLICANT: SCHUCH, WOLFGANG W.
; TITLE OF INVENTION: DNA, CONSTRUCTS, CELLS AND PLANTS
; TITLE OF INVENTION: DERIVED THEREFROM
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: CUSHMAN DAREY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/300,582
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/859,523
; FILING DATE: 12-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 3893/95874
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-300-582-2
;
; Alignment Scores:
; Pred. No.: 2,94e-193 Length: 1646
; Score: 1732.00 Matches: 346
; Percent Similarity: 85.68% Conservative: 31
; Best Local Similarity: 78.64% Mismatches: 36
; Query Match: 76.30% Indels: 27
; Gaps: 4
;
; US-09-847-081B-2 (1-440) x US-08-300-582-2 (1-1646)
;
; Qy 1 MetSerMetSerValAlaLeuLeuTyrValValSerProThrSerGluValSerAsnGly 20
; Db 195 CTCGAATGCTGTTGCTTGTATAGGTTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 251
;
; Qy 21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40
; Db 252 ACAAGTTTACGAAATAGTCGCGGAGGAGAAACCGTTTTTTATGATATGCTCTCTCTCT 302
;
; Qy 41 AlaArgAspArgAsnLeuMetTyrPasnGlyArgIleLysLysGlyGlyArgGlnArgTyr 60
; Db 303 ---AGCATAGAGAAATTTGGTGTCCAAATGAGAGAAATCAATAGAGGT----- 344
;
; Qy 61 AsnPheGlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGlu 80
; Db 345 -----GGTGAAGAGCAAACTAAT 362
;
; Qy 81 LysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThr 100
; Db 363 AATGACGGAATTTTCTGACGCTGCTATTTTGGCTACTCATCTGAGAGAGCGAGC 422
;
; Qy 101 ValSerSerGluLysValTyrAspValValLeuLysGlnAlaLeuValLysArg 120
; Db 1201 ValSerSerGluLysValTyrAspValValLeuLysGlnAlaLeuValLysArg 120

```

Db 423 ATGACATCGGAACAGATGCTATGATGCTTTTGAAGCAGGACCTTGTGAAAGG 482  
 QY 121 GlnLeuArgSerThrAspAspLeuGluValLysProAspIleValProGluAsnLeu 140  
 Db 483 CAAGTGAATCTTACCAATAGATGTAAGGAAAGCCGGAATATCTTATCCGGGAATTGG 542  
 QY 141 GlyLeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThr 160  
 Db 543 GCGTGTGATGAGAACATATGATGATGCTGAGTATGTCAGATATGCAAGAGG 602  
 QY 161 PheTyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrrPalaIleTyr 180  
 Db 603 TTTAACTTGAAGAACTATGCTAAATGACTCCGAGAGAAAGAGGCTATCGGCAATATAT 662  
 QY 181 ValTrrPheArgArgThrAspGluLeuValAspGluProAsnAlaSerHisIleThrPro 200  
 Db 663 GATATGTTGCAAGAAACAGATGAACTTGTGATGCGCCAAAGCCATCATATATACCCG 722  
 QY 201 GlnAlaLeuAspArgTrrPgluThrArgLeuGluAspIlePheSerGlyArgProPheAsp 220  
 Db 723 GCAGCCTTATGATGCTGAGGAAATAGGCTAAGAAATGTTTCAATGGCGGCATTTGAC 782  
 QY 221 MetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPhe 240  
 Db 783 ATGCTGATGCTGCTTGTTCGACATGACATTTCACTTCACTTCAATATTCAGCCATTC 842  
 QY 241 ArgAspMetIleGluGlyMetArgMetAspLeuTrrLysSerArgTyrLysThrPheAsp 260  
 Db 843 AGAGATATATTCAGAAATGCTATGCTATGACCTTGAAGAAATCGAGATACAAATCTTCAC 902  
 QY 261 GlnLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProVal 280  
 Db 903 GAACATATACCTTATATGTTATATGTTGCTGAGTACCGTGGGTGATGAGTGTTCACAT 962  
 QY 281 MetGlyIleAlaProGluSerLysAlaIleThrGluSerValTyrAsnAlaAlaLeuAla 300  
 Db 963 ATGGGATGCGCCCTGAATCAAGCAACAGAAAGGCTATATATCTGCTTGGCT 1022  
 QY 301 LeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArg 320  
 Db 1023 CTGGGGATCGCAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1082  
 QY 321 GlyArgValTyrLeuProGluAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIle 340  
 Db 1083 GGAAGAGTCTACTGCTGCTCAAGATGAATTAGCAAGAGGCTATATCCATGAAAGTATAT 1142  
 QY 341 PheAlaGlyArgValThrAspLysTrrPheAsnMetLysGlnIleGlnArgAla 360  
 Db 1143 TTTGCTGGAAGGCTGACCGATTAATGAGATCTTATATGAAACAAATATCATAGGCGCA 1202  
 QY 361 ArgLysPhePheAspGluSerGlyLysGlyValThrGluLeuAspSerLysArgTrrPhe 380  
 Db 1203 AGAAATCTTGTATGAGGAGGAGAAAGGCTGACCAATGAGCTCACTCATGATATATTC 1262  
 QY 381 ProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnArg 400  
 Db 1263 CCGTATGAGGCACTTGTGCTTGTATCCGCAAAATCTGATGAGTTGAGCCATATGAC 1322  
 QY 401 TyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuPro 420  
 Db 1323 TCAACCACTTCAACAAAGAGAGCATATGAGCAAAATCAAGCAATGTTGATTCATTAC 1382  
 QY 420 GlnLeuAlaTyrAlaLysSerLeuValProProAsnArgThrSerSerProLeuAlaLys 439  
 Db 1383 TATTCATATGCAAAATCTTGTGCTCTCT-----ACAAAATGCTCTCTTCAAA 1434  
 RESULT 9  
 US-08-579-667-9  
 : Sequence 9, Application US/08579667  
 : Patent No. 5705624  
 : GENERAL INFORMATION:  
 : APPLICANT: Fitzmaurice, Wayne P.  
 : APPLICANT: Hellmann, Gary M.

: APPLICANT: Grill, Laurence K.  
 : APPLICANT: Kumagai, Monto H.  
 : APPLICANT: Della-Cioppa, Guy R.  
 : TITLE OF INVENTION: DNA SEQUENCES ENCODING ENZYMES USEFUL IN  
 : TITLE OF INVENTION: PHYTOENE BIOSYNTHESIS  
 : NUMBER OF SEQUENCES: 19  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESS: Virginia C. Bennett  
 : STREET: 1211 East Morehead Street, PO Drawer 34009  
 : CITY: Charlotte  
 : STATE: No. 5705624th Carolina  
 : COUNTRY: USA  
 : ZIP: 28234  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent In Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/579,667  
 : FILING DATE:  
 : CLASSIFICATION: 435  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Bennett, Virginia C.  
 : REGISTRATION NUMBER: 37,092  
 : REFERENCE/DOCKET NUMBER: 627-196  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 919-420-2200  
 : TELEFAX: 919-861-3175  
 : INFORMATION FOR SEQ ID NO: 9:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 749 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: cDNA  
 : US-08-579-667-9  
 : Alignment Scores:  
 : Pred. No.: 2,15e-42 Length: 749  
 : Score: 442.00 Matches: 98  
 : Percent Similarity: 73.83% Conservative: 12  
 : Best Local Similarity: 65.77% Mismatches: 11  
 : Query Match: 19.47% Indels: 28  
 : DB: 1 Gaps: 4  
 : US-09-847-081b-2 (1-440) x US-08-579-667-9 (1-749)  
 QY 3 MetSerValAlaLeuLeuTrrPheValSerProThrSerGluValSerAsnGlyThrGly 22  
 Db 382 ATGCTGTTGCTTGTATGCTTGTATGCTTGTATGCTTGTATGCTTGTATGCTTGTATGCT 438  
 QY 23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42  
 Db 439 TTTCTTGATTCAGAAAGGAGGAAACCGGCTTTTGAATCTGCTG-----AGG 486  
 QY 43 AspArgAsnLeuMetTrrPheAsnGlyArgIleLysGlyGlyValArgGlnAlaArgTrrPhe 62  
 Db 487 CATAGAAATTTAGTGTGCATATGAGAAATCAAAAGAGGTGTAAACAAAGGTGCAATTTT 546  
 QY 63 GlySerLeuIleAlaAspProArgTrrSerCysLeuGlyGlySerArgThrGluLysGly 82  
 Db 547 GGT----- 549  
 QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMet---ThrVal 101  
 Db 550 -----TCCGTCAGGCTGCTGCATGCTGCTACACCAACGAGAAATGCGCAATG 600  
 QY 102 SerSerGluLysValTrrPheValValLeuLysGlnAlaAlaLeuValLysArgGln 121  
 Db 601 ACATCAGAAACAGAGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
 QY 122 LeuArgSerThrAspAspLeuGluValLysProAspIleValProGlyAsnLeuGly 141

Db 661 CTGAGATCTACTGATGATTTAGAACTGAAGCGGAGATCCCTCCCGGGAATTTGAGC 720  
Qy 142 LeuLeuSerGluAlaTyrAspArgCys 150  
Db 721 TTGTTAAGTGAAGCATATGATGATGCTGT 747

RESULT 10  
US-08-095-726-5  
; Sequence 5, Application US/08095726  
; Patent No. 5530189  
; GENERAL INFORMATION:  
; APPLICANT: Auslich, Rodney L  
; APPLICANT: Brinkhaus, Friedhelm L  
; APPLICANT: Mukharji, Indrani  
; APPLICANT: Proffitt, John H  
; APPLICANT: Yarger, James G  
; APPLICANT: Yen, Huel-Che B  
; TITLE OF INVENTION: Beta-Carotene Biosynthesis in  
; TITLE OF INVENTION: Genetically Engineered Hosts  
; NUMBER OF SEQUENCES: 79  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept  
; STREET: 200 E Randolph St  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60680-0703  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.24  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/095,726  
; FILING DATE: 21-JUL-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/785,566  
; FILING DATE: 30-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Galloway, No. 5530189val B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 3128567180  
; TELEFAX: 3128564972  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1198 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-095-726-5

Alignment Scores:  
Pred. No.: 6.55e-33 Length: 1198  
Score: 364.00 Matches: 95  
Percent Similarity: 50.53% Conservative: 49  
Best Local Similarity: 33.33% Mismatches: 124  
Query Match: 16.04% Indels: 18  
DB: 1 Gaps: 4

US-09-847-081b-2 (1-440) x US-08-095-726-5 (1-1198)

Qy 148 AsparGysGlyGluValCysAlaGluTyrAlaLysThrPheTyrLeuGlyThrLysLeu 167  
Db 37 GACCAAGCGACGACGACCATGCGCAACGCGCTCGAAGATTTCGACACCGCTGCAAGCTG 96  
Qy 168 MetThrProGluArgArgAlaIleTyrPalaIleTyrValTyrCysArgArgThrAsp 187  
Db 97 TTCGACCGCGGACCGCGCTGAGCGTGTACACCTGATGCGCGACGCTGCAAT 156  
Qy 188 GluLeuValAspGlyProAsn-----AlaSerHisIleThr 199

Db 157 GACCTCATTCGACGACACGACCGCTTCGCCAGGAGCGCGGAGAGAGAGGCC 216  
Qy 200 ProGlnAlaLeuAspArgTyrPglThrArgLeuGlnAspIlePheSerGlyArgProPhe 219  
Db 217 ACCGAGCGCTGCGCGCGCTGCGGCACTGACCTGCGCGGCTTTGAAAGGCGCGAATG 276  
Qy 220 ---AspMetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGln 238  
Db 277 CAGGATCCGGCTTCGCTGCTTCAGAGAGTGCGCGTGAACCCAGGATATGACCCCGC 336  
Qy 239 ProPheArgAspMetIleGluGlyMetArgMetAspLeuTyrPlySerArgTyrLysThr 258  
Db 337 ATGCGGCTGCATCACTTCGACGCTTCGATGACGTGCGTGCAGACCCGCTATGTCACC 396  
Qy 259 PheAspGluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerVal 278  
Db 397 TTGAGGATACGCTGCGCTGCTGCTATCAGTGGCGGCGTGTGGCTGATATGAGCC 456  
Qy 279 ProValMetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAla 298  
Db 457 AGGGTATGGCGCG-----CGGATGAGCGGCTGCTGATCGCGCC 498  
Qy 299 LeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAla 318  
Db 499 TGCATCTGGGCGTGGCTTCACGCTGACGATATGCGCGGATATATTGACGATGCG 558  
Qy 319 ArgArgGlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGlu 338  
Db 559 GCTATTGACCGCTGCTATCTGCCCGGAGGCTGAGAGATGCGGCGCTGCGCCGAG 618  
Qy 339 AspIlePheAla-GlyArgValThrAspLysTyrArgAsnMetLysGlnIleGln 358  
Db 619 AACATACCGCGCGGAGAAATCGCCCGCGCTGCGCGG--TGG--CGAGGCTTATTGA 674  
Qy 358 nArgAlaArgLysPheAspGluSerGlyLysGlyValThrGluLeuAspSerAlaSe 378  
Db 675 TGCCGACAGCGCTACATCTCTCCACAGCGCGGCTACACGATGCGCGCGCTC 734  
Qy 378 TArgTyrProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGlnIleGln 398  
Db 735 CGCGTGGCGGATCGCACCGCGCGCGACGCTTACCGGAGATCGGTAAAGTAAAGC 794  
Qy 398 AsnAspTyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuTh 418  
Db 795 GCGCGAGGACGCGCTGGATCGCCCGCACACACGACGACAAAGTAAAAATTGCCAT 854  
Qy 418 rLeuProIleAla 422  
Db 855 GCTGATGCGGCA 867

RESULT 11  
US-08-096-043-5  
; Sequence 5, Application US/08096043  
; Patent No. 5530189  
; GENERAL INFORMATION:  
; APPLICANT: Auslich, Rodney L  
; APPLICANT: Brinkhaus, Friedhelm L  
; APPLICANT: Mukharji, Indrani  
; APPLICANT: Proffitt, John H  
; APPLICANT: Yarger, James G  
; APPLICANT: Yen, Huel-Che B  
; TITLE OF INVENTION: Lycopen Biosynthesis in  
; TITLE OF INVENTION: Genetically Engineered Hosts  
; NUMBER OF SEQUENCES: 70  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept  
; STREET: 200 E Randolph St  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60680-0703  
; COMPUTER READABLE FORM:







```

Db 97 TTCCAGCCCGGACACCGCGCGTACGGCTGTAGTCTCTACACCTGTGGCGCCACTCGAT 156
Qy 188 GlnueValaspGlyProasn-----AlaserHisIleThr 199
Db 157 GACGTCATTTAGACACACGACCGACTTCGCCAGCAGAGCCGCGCGCAGAGAGAGCC 216
Qy 200 ProGlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPhe 219
Db 217 ACCCAGCGCCCTGGCCCGGCTGGCGCAGCTGACCTCGGCGCGCTTTGAAAGGCCAGATG 276
Qy 220 --AspMetLeuAspAlaLeuSerAspThrValSerArgPheProValAspIleGln 238
Db 277 CAGATATCCGCGCTTCGCTGCTCTTTGAGAGAGTGCGCTGACCCACGAGTATTACGCCCGC 336
Qy 239 ProPheArgAspMetIleGluGlyMetArgMetAspLeuTrpIysSerArgTrpThr 258
Db 337 ATGGCGCTCGATCACCCTCGACGCGCTTTGGCGATGAGTGAGTGCGCTTCACCCGCTGTGCACC 396
Qy 259 PheAspGlnLeuTrpLeuTrpCysTrpTrpValAlaGlyThrValGlyLeuMetSerVal 278
Db 397 TTTCAGGATACCGTCCGCTGCTACTGCTATCAGTGCGCGCGTGCGTGGCTGTGATGATGCC 456
Qy 279 ProValMetGlyIleAlaProGlnSerIysAlaThrThrGlnSerAlaTrpAsnAlaIa 298
Db 457 AGCGGATGGCGGTG-----CGGAGATGAGCGCGGTCTCGATCCGCC 498
Qy 299 LeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAla 318
Db 499 TGCGATCTGGGGCTGGCTTCAGCTGACGAAATATAGCCCGGAGATATTATTCAGATGCG 558
Qy 319 ArgArgGlyArgValTrpLeuProGlnAspGlnLeuAlaGlnAlaGlyLeuSerAspGln 338
Db 559 GCTATTGACCGCTGCTATTCGCCCGCAGATGCGTGCAGAGATCCGAGCTGGCCCGAG 618
Qy 339 AspIlePheAla-GlyArgValThrAspTrpArgAsnPheMetIysGlnIleG 358
Db 619 AACATATGCCCGCGCGGAGATGCGCCCGCGCTGCCCG-TGS---CGAGCGTTATTGA 674
Qy 358 nArgAlaArgIysPhePheAspGlnSerGluIysGlyValThrGlnLeuAspSerAlaSe 378
Db 675 TGCCGCAAGCGCTACTACATCTCTCCAGCGCGGCGTACAGATCTGGCGCGCGCTC 724
Qy 378 ArgTrpProValLeuThrAlaLeuLeuLeuTrpArgIysIleLeuAspGlnIleGluAl 398
Db 735 CGCGTGGCGATCGGCACCGCCGCGCAGCGCTTACCGGAGATGGATTAAAGGTAAAGC 794
Qy 398 aAsnAspTrpAsnAsnPheThrArgArgAlaTrpValSerIysProIysIysLeuTh 418
Db 795 GCGCGGAGAGCGCGCTGGGATGCGCGCAGCACACAGCAAGGTGAATAAATTCGCAT 854
Qy 418 rLeuProIleAla 422
Db 855 GCTGATGGCGCA 867

RESULT 14
US-08-331-004A-1
; Sequence 1, Application US/08331004A
; Patent No. 5618988
; GENERAL INFORMATION:
; APPLICANT: Hauptmann, Randal
; APPLICANT: Eschenfeldt, William H
; APPLICANT: English, Jaml
; APPLICANT: Brinkhaus, Friedhelm I
; TITLE OF INVENTION: Enhanced Carotenoid Accumulation
; TITLE OF INVENTION: in Storage Organs of Genetically
; TITLE OF INVENTION: Engineered Plants
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corporation, Law Dept
; STREET: 55 Shuman Boulevard, Suite 600
; CITY: Naperville
; STATE: IL
; COUNTRY: USA

```

[illegible]

Qy 339 Aspllepheala-glyargvalthrAspLysThrPargAsnPhenetylsylsGlnIleG1 358  
 Db 606 AACATATGCCGCGCGGAGATCGGGCGCGCTGCGGGGTGGCG-GAGCGGCTTATTGA 664  
 Qy 358 nargAlAArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSeraLase 378  
 Db 665 TGCCGACAGCCGCTACTCTCTCCACGCGCGGTACACGATCTCCGCGCGCTG 724  
 Qy 378 rArgTrpProValLeuThrAlaLeuLeuLeuThrArgLysIleLeuAspGluIleGluA1 398  
 Db 725 GCGCTGGCGCATCGCCACCGCCGCGACGCTCTACCGGAGATCGGTATTAAAGTAAAGC 784  
 Qy 398 aAsnAspTyraAsnAsnPhenThrArgArgAlaTyraValSerLysProLysLysLeuLeuTh 418  
 Db 785 GCGCGAGAGCAGCGCGCTGAGATCGCGCGACACACGAAAGGTGAAAAATTGCCAT 844  
 Qy 418 rLeuProIleAla 422  
 Db 845 GCTGATGCGGCA 857

## RESULT 15

PCT-US95-13937A-1  
 : Sequence 1, Application PC/TUS9513937A  
 : GENERAL INFORMATION:  
 : APPLICANT: Hauptmann, Randal  
 : APPLICANT: Eschenfeldt, William H  
 : APPLICANT: English, Jami  
 : APPLICANT: Brinkhaus, Friedhelm L  
 : TITLE OF INVENTION: Enhanced Carotenoid Accumulation  
 : TITLE OF INVENTION: In Storage Organs of Genetically  
 : NUMBER OF SEQUENCES: 9  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Amoco Corporation, Law Dept  
 : STREET: 55 Shuman Boulevard, Suite 600  
 : CITY: Naperville  
 : STATE: IL  
 : COUNTRY: USA  
 : ZIP: 60563-8437  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent Release #1.24  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: PCT/US95/13937A  
 : FILING DATE:  
 : CLASSIFICATION:  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Galloway, Norval B  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 7087172447  
 : TELEFAX: 7087172430  
 : INFORMATION FOR SEQ ID NO: 1:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 1083 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: DNA (genomic)  
 : PCT-US95-13937A-1

## Alignment Scores:

Pred. No.: 1.09e-32 Length: 1083  
 Score: 361.50 Matches: 95  
 Percent Similarity: 50.18% Conservative: 48  
 Best Local Similarity: 33.33% Mismatches: 126  
 Query Match: 15.93% Indels: 17  
 DB: 5 Gaps: 3

US-09-847-081B-2 (1-440) x PCT-US95-13937A-1 (1-1083)

Qy 148 AspArgCysGlyLeuValCysAlaGluTyraIleuThrPheTyreGluGlyThrLysLeu 167  
 Db 24 GACCACGCCACGACGACATAGCCCAACGCGCTCGAAAAAGTTTCCACCGCTGGCAAGCTT 83  
 Qy 168 MetThrProGluArgAlaArgAlaIleThrAlaIleTyraValThrProCysArgAlaThrAsp 187  
 Db 84 TTCGACCCGGCCACCCGCGCTAGGCTGATGCTCTACACCTGGTGGCCACTGGCAT 143  
 Qy 188 GluLeuValAspGlyProAsn-----AlaSerHlaIleThr 199  
 Db 144 GACCTCATTAACACACGACGACCAACGCGCTTCCAGAGAGGCGCGGAGAGAGAGGCC 203  
 Qy 200 ProGlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPhe 219  
 Db 204 ACCCAGCGCTGGCGCGCGCTGCGCACGCTGACCTGGCGCGCTTTGAAGGGGCGAGATG 263  
 Qy 220 ---AspMetLeuAspAlaIleuSerAspThrValSerArgPheProValAspIleGln 238  
 Db 264 CAGGATCCGCGCTTGGCTGCTTTCAGAGAGTGGCGCTGACCCAGGTATTACGCCCCGC 323  
 Qy 239 ProPheArgAspMetIleGluGlyMetArgMetAspLeuThrLysSerArgTyraThr 258  
 Db 324 ATGGCGCTCGATACCTCGACGCGCTTGGCATGAGACGTGCTGACAGCCCTGATATGTCACC 383  
 Qy 259 PheAspGluLeuTyreTyraValAlaGlyThrValGlyLeuMetSerVal 278  
 Db 384 TTTCAGGATACGCTGCGCTACTGCTACTGATCAGTGGCGGCGTGGTGGCTGATGATGGCC 443  
 Qy 279 ProValMetGlyIleAlaProGluSerLysAlaThrThrLysSerValTyraAsnAla 298  
 Db 444 AGGCTGATGGCGG-----CGGATGACGCGGTCTGGATCGCGCC 485  
 Qy 299 LeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAla 318  
 Db 486 TGGCATTTGGGGGCTGGCTTCCAGCTGACGATATGCCCGGATATTATTAAGATGCGG 545  
 Qy 319 ArgArgGlyArgValTyreProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGlu 338  
 Db 546 GCTATTGACCGCTACTCTGCGCGCGAGTGGCTGACGAGATCGCGGCTGACCCCGAG 605  
 Qy 339 Aspllepheala-glyargvalthrAspLysThrPargAsnPhenetylsylsGlnIleG1 358  
 Db 606 AACATATGCCGCGCGGAGATCGGGCGCGCTGCGGGGTGGCG-GAGCGGCTTATTGA 664  
 Qy 358 nargAlAArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSeraLase 378  
 Db 665 TGCCGACAGCCGCTACTCTCTCCACGCGCGGTACACGATCTCCGCGCGCTG 724  
 Qy 378 rArgTrpProValLeuThrAlaLeuLeuLeuThrArgLysIleLeuAspGluIleGluA1 398  
 Db 725 GCGCTGGCGCATCGCCACCGCCGCGACGCTCTACCGGAGATCGGTATTAAAGTAAAGC 784  
 Qy 398 aAsnAspTyraAsnAsnPhenThrArgArgAlaTyraValSerLysProLysLysLeuLeuTh 418  
 Db 785 GCGCGAGAGCAGCGCGCTGAGATCGCGCGACACACGAAAGGTGAAAAATTGCCAT 844  
 Qy 418 rLeuProIleAla 422  
 Db 845 GCTGATGCGGCA 857

Search completed: January 16, 2004, 08:48:56  
 Job time : 104 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 16, 2004, 08:00:33 ; Search time 444 Seconds  
(without alignments)  
3493.009 Million cell updates/sec

Title: US-09-847-081B-2

Perfect score: 2270  
Sequence: 1 MSMSVALLMVVSPTESEVNSG.....IYAKSLVPPNRTSSPLAKT 440

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n -DEV=xlh  
-O=/cg2\_1/USPTO/US09847081/runat\_15012004\_154339\_1958/app\_query.fasta\_1.583  
-DB=Published Applications NA -QPM1=fastap -SUFF1=trpb -MINMATCH=0.1  
-LOOPEL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=plc -NORM=ext -HEATSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09847081 @CGN 1.1 271 @runat\_15012004\_154339\_1958  
-NCPU=6 -ICPU=3 -NO MAP -LARS=JURY -NEG\_SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cg2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cg2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cg2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cg2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cg2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cg2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cg2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cg2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cg2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cg2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
11: /cg2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
12: /cg2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cg2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
14: /cg2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
15: /cg2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cg2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
17: /cg2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
18: /cg2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
------------	-------	-------------	-----------	----	-------------

1	2270	100.0	1728	10	US-09-847-081B-1	Sequence 1, Appli
2	1801	79.3	1712	10	US-09-847-081B-3	Sequence 3, Appli
3	1737	76.5	1239	9	US-09-371-307-75	Sequence 75, Appli
4	1737	76.5	1239	12	US-10-401-321-75	Sequence 75, Appli
5	1542	67.9	1269	10	US-09-938-842A-729	Sequence 729, App
6	468	20.6	684	13	US-10-149-759-53	Sequence 53, Appli
7	372	16.4	1026	15	US-10-156-761-1013	Sequence 1013, Ap
8	372	16.4	9025608	15	US-10-156-761-1	Sequence 1, Appli
9	368	16.2	948	15	US-10-156-761-1642	Sequence 1642, Ap
10	368	16.2	9025608	15	US-10-156-761-1	Sequence 1, Appli
11	359.5	15.8	945	13	US-10-128-713A-15	Sequence 15, Appli
12	350.5	15.4	1232	14	US-10-041-472-1	Sequence 1, Appli
13	336.5	14.8	891	11	US-09-941-947A-33	Sequence 33, Appli
14	336.5	14.8	891	13	US-10-218-118-9	Sequence 9, Appli
15	327.5	14.4	1008	15	US-10-166-037-4	Sequence 4, Appli
16	326	14.4	8625	11	US-09-920-923-1	Sequence 1, Appli
17	326	14.4	11233	11	US-09-920-923-27	Sequence 27, Appli
18	324	14.3	908	10	US-09-547-267-4	Sequence 4, Appli
19	306	13.5	1020	11	US-09-918-740-71	Sequence 71, Appli
20	306	13.5	7252	11	US-09-918-740-71	Sequence 73, Appli
21	306	13.5	7252	11	US-09-918-740-75	Sequence 75, Appli
22	293	12.9	536165	11	US-09-939-964-1	Sequence 1, Appli
23	273.5	12.0	2174	8	US-08-781-986A-202	Sequence 202, App
24	272.5	12.0	864	13	US-10-358-917-9	Sequence 9, Appli
25	230	10.1	3807	12	US-10-438-784-2	Sequence 2, Appli
26	230	10.1	5898	12	US-10-438-784-1	Sequence 1, Appli
27	221	9.7	825	12	US-10-438-784-6	Sequence 6, Appli
28	204	9.0	912	10	US-09-738-626-695	Sequence 695, App
29	204	9.0	3309400	10	US-09-738-626-1	Sequence 1, Appli
30	198	8.7	876	10	US-09-738-626-2673	Sequence 2673, Ap
31	197.5	8.7	905	15	US-10-156-761-1643	Sequence 1643, Ap
32	170	7.5	405	13	US-10-259-678-345	Sequence 345, App
33	167.5	7.4	1092	13	US-10-358-917-1	Sequence 1, Appli
34	161	7.1	1173	12	US-10-369-493-42261	Sequence 42261, A
35	145.5	6.4	408	15	US-10-166-037-8	Sequence 8, Appli
36	141	6.2	1347	13	US-10-032-585-6355	Sequence 6355, Ap
37	139.5	6.1	1335	12	US-10-369-493-45750	Sequence 45750, A
38	138.5	6.1	1587	15	US-10-128-714-7505	Sequence 7505, Ap
39	135.5	6.0	1563	15	US-10-128-714-2505	Sequence 2505, Ap
40	135.5	6.0	1665	15	US-10-128-714-1505	Sequence 1505, Ap
41	135.5	6.0	1730	15	US-10-128-714-5505	Sequence 5505, Ap
42	135.5	6.0	3665	15	US-10-128-714-5505	Sequence 5505, Ap
43	135.5	6.0	3730	15	US-10-128-714-5505	Sequence 5505, Ap
44	126	5.6	2097	13	US-10-240-965-172	Sequence 172, App
45	123	5.4	984	15	US-10-156-761-3904	Sequence 3904, Ap

## ALIGNMENTS

RESULT 1  
US-09-847-081B-1  
Sequence 1, Application US/09847081B  
Patent No. US20020128464A1  
GENERAL INFORMATION:  
APPLICANT: BAYER AG  
TITLE OF INVENTION: DNA encoding the tobacco phytoene synthase  
FILE REFERENCE: Le A 34 326  
CURRENT APPLICATION NUMBER: US/09/847,081B  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 1728  
TYPE: DNA  
ORGANISM: Nicotiana tabacum  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (244)..(1566)  
US-09-847-081B-1  
Alignment Scores: 1.17e-274 Length: 1728  
Pred. No.: 2270.00 Matches: 440  
Score:

Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 100.00%  
 DB: 10  
 Gaps: 0

US-09-847-081B-2 (1-440) x US-09-847-081B-1 (1-1728)

1 MetSerMetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGly 20  
 244 ATGAGCATGTCCTGTTGCTTGTGGGTGTTGTTCCACCTCCGAGGTCTCGAATGG 303  
 21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40  
 304 ACAGATGTTGGATTGATTCAGTCCGAGAGAAACCCGCTTGTGATCATCCAGGTGCTCTA 363  
 41 AlaArgAspArgAsnLeuMetTrpAsnGlyArgGlyLeuValGlyValArgGlyArgTrp 60  
 364 GCTCGAGATGAGAAATTTATGATGATGAGATGAGAAATCAAGAAAGTGGAGACAAAGGTGG 423  
 61 AsnPheGlySerLeuLeuAlaAspProArgTrpSerCysLeuGlyGlySerArgTrpGlu 80  
 424 AATTGGCTCTTATTCCTGATGATCCAGATATTCCTGCTGGTGATCAAGAACTGAA 483  
 81 LysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThr 100  
 484 AAGGAAACACCTTCTCTGTACAGTCCAGTTGGTGGCTAGCCGAGTGGAGAAATGACT 543  
 101 ValSerSerGluValGlyValValTrpAspValValLeuValGlnAlaLeuValLeuArg 120  
 544 GTGTCATCAGAGAAAGAGGTATGATGTGATTTAAAGCGGCGCTTGTAGTGAGAGG 603  
 121 GlnLeuArgSerThrAspAspLeuGluValLysProAspIleValProGlyAsnLeu 140  
 604 CAGCTGAGATCTTACCGATGATTTAGAAAGTGAAGCCGATTTGTTCCAGGGAATTTG 663  
 141 GlyLeuLeuSerGluAlaTrpAspArgCysGlyGluValCysAlaGlyTrpAlaLysThr 160  
 664 GCGTTGTTGAGTGAAGCATATGATCGTTGGCGAAGATGAGCAGATGATGCAAGACA 723  
 161 PheTrpLeuGlyThrLysLeuMetThrProGluArgGlyAlaIleTrpAlaIleTrp 180  
 724 TTTTACTTAGAGAACCAACTATGATGCCAGAGAAAGAGCATATCTGGCAATATAT 783  
 181 ValTrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrPro 200  
 784 GTGTCATCAGAGAAAGAGGTATGATGTGATGATGATGATGATGATGATGATGATGAT 843  
 201 GlnAlaLeuAspArgTrpGluThrArgLeuGluValAspIlePheSerGlyArgProPheAsp 220  
 844 CAAGCTTAGATGAGTGGAGAACCAAGCTGGAAGATATTTTCAGTGGCGGCAATTCAT 903  
 221 MetLeuAspAlaAlaLeuSerThrPheValSerArgPheProValAspIleGlnProPhe 240  
 904 ATGCTTAGATGAGTGGAGAACCAAGCTGATCTCCCAATTTCCGTGATATTCAGCCATTC 963  
 241 ArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTrpLysThrPheAsp 260  
 964 AGAGATGATGATGAGTGGAGAACCAAGCTGATGATGATGATGATGATGATGATGATGAT 1023  
 261 GluLeuTrpLeuTrpCysTrpTrpValAlaGlyThrValGlyLeuMetSerValProVal 280  
 1024 GAGCATATCTCTATATGTTATGATGTTGCTGTTATGATGATGATGATGATGATGATGAT 1083  
 281 MetGlyIleAlaProGluSerLysAlaThrThrGluSerValTrpAsnAlaAlaLeuAla 300  
 1084 ATGGTATTTGACCTCGAATCAAGGCAACACAGAGAGTATATTAATGCTGCTTGGCT 1143  
 301 LeuGlyLeuAlaAsnGlyLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArg 320  
 1144 TTAGGCTTGGCAATCAACTAACATATACTCAAGATGATGAGAGAGTGGCAAGAGA 1203  
 321 GlyArgValTrpLeuProGlnAspGluLeuValAlaGlyLeuSerAspGluAspIle 340

1204 GGAAGATATACTTGGCTCAAGATGAATTAGCAAGGCGAGGCTCTCCAGCAAGACATA 1263  
 341 PheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArgAla 360  
 1264 TTTCCTGAGAGAGTGAATGATGATGAGAGAGAACTTTATGAGAAACAAATTCAGAGGCG 1323  
 361 ArgLysPhePheAspGluSerGluValValThrGluLeuAspSerAlaSerArgTrp 380  
 1324 AGGAATATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1383  
 381 ProValLeuThrAlaLeuLeuLeuTrpArgLysIleLeuAspGluIleGlnAlaAsnAsp 400  
 1384 CCGTGTGTTACAGCGCTGCTGTGTATGCGCAAGATATTTGACGAGATTGAAGCCAAAGC 1443  
 401 TyrAsnAsnPheThrArgAlaTrpValSerLysProLysLysLeuLeuThrLeuPro 420  
 1444 TACACAACTTTCACAGAGAGGCTTATGTTATGCAAGCCAAAGAGCTTTCACCTTCCC 1503  
 421 IleAlaTrpAlaLysSerLeuValProProAsnArgThrSerSerProLeuAlaLysThr 440  
 1504 ATTGCTATGCAAAATCTCTGTGCCCCCTAATGAACTTCCTCTCAGTGAAGAGACA 1563

# RESULT 2

US-09-847-081B-3  
 ; Sequence 3, Application US/09847081B  
 ; Patent NO. US20020128464A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BAYER AG  
 ; TITLE OF INVENTION: DNA encoding the tobacco phytoene synthase  
 ; FILE REFERENCE: Le A 34 326  
 ; CURRENT APPLICATION NUMBER: US/09/847,081B  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 3  
 ; LENGTH: 1712  
 ; TYPE: DNA  
 ; ORGANISM: Nicotiana tabacum  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (333)..(1565)  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: 135..139  
 ; OTHER INFORMATION: Xaa is unknown or other  
 ; NAME/KEY: unsure  
 ; LOCATION: 51  
 ; OTHER INFORMATION: n can be any nucleotide  
 US-09-847-081B-3

# Alignment Scores:

Pred. No.: 1.04e-215 Length: 1712  
 Score: 1801.00 Matches: 357  
 Percent Similarity: 87.93% Conservative: 29  
 Best Local Similarity: 81.32% Mismatches: 24  
 Query Match: 79.34% Indels: 29  
 DB: 10 Gaps: 4

US-09-847-081B-2 (1-440) x US-09-847-081B-3 (1-1712)

3 MetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThrGly 22  
 333 ATGCTGTGCTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 389  
 23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42  
 390 TTCTTGATTCAGTCCGAGAGGAGAAACCGGTTTATTCATTCGTG-----AGG 437  
 43 AspArgAsnLeuMetTrpAsnGlyArgIleLysLysGlyValArgGlnAlaGlyTrpAsnPhe 62  
 438 CATAGCAATTTAGTGTCAATGACAGAGAAACAGAGAGGTGTGAAACAAAGGTGGAATTTT 497

QY 63 GlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluLeuGly 82  
 Db 498 GGT----- 500  
 QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMet---ThrVal 101  
 Db 501 -----TCGTGAAGCTCTGCTATGGCTGACACGGCGGAGAAATGGCGACGATG 551  
 QY 102 SerSerGluLeuValTyrAspValValLeuValGlnAlaAlaLeuValTyrAspGln 121  
 Db 552 ACATCAGAAACAGATGTTATGATGTTGTTTAAACAAGACGCTTTGTAAGAGGACG 611  
 QY 122 LeuArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeuGly 141  
 Db 612 TTGAATCTGCTGATGATTTAGAGTGAAGCGAGATCCCTCCCGGGAATTTGAGC 671  
 QY 142 LeuLeuSerGluValTyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThrPhe 161  
 Db 672 TTGTGAGTGAAGCATGATGATGCTGAGTGAAGTATGTCAGAGTATGCACAAAGCATT 731  
 QY 162 TyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTyrAlaIleTyrVal 181  
 Db 732 TACTHGGAGACCATGTATATGACTCCAGAGAGAAAGGCTATTGGCATAATATGTG 791  
 QY 182 TrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGln 201  
 Db 792 TGGTCAGAGAGAAACAGATGACTTGTGATGCGCCAAAGCATCATATTAACCCCA 851  
 QY 202 AlaLeuAspArgTrpGluThrArgLeuGluValIlePheSerGlyArgProPheAspMet 221  
 Db 852 GCCTTAGAGATGGTGGAAAGACCGGCTTGAAGATGTTTCAAGGCGACCATTTGATATG 911  
 QY 222 LeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArg 241  
 Db 912 CTCGATGCTGCTTTCGATTCGATCTGTTCCAGTTTCCAGTTGATATTCAGCCGTTCA 971  
 QY 242 AspMetIleGluGlyMetArgMetAspLeuTyrLysSerArgTyrLysThrPheAspGlu 261  
 Db 972 GATATGATTTGAAGAGATGGGTATGACTTGAAGAGTCAAGATATAGAATCTTATGAG 1031  
 QY 262 LeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMet 281  
 Db 1032 CTTTACCTTATGTTATTAACGTCCTGCTGATGAGTGGTGTATGATGATGATTCATTA 1091  
 QY 282 GlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeu 301  
 Db 1092 GGTATTCGACCTGATTCAAAGCAACACAGAGCGTATATATGACCTTTGGCTTTA 1151  
 QY 302 GlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGly 321  
 Db 1152 GGAATCGCAATCACTAAGCAACATTACTCAGAGATGTTGAGAGATGCCAGAAAGAGA 1211  
 QY 322 ArgValIleTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePhe 341  
 Db 1212 AGAGCTCACTTAACCTCAAGATGAATTAAGACAGGAGGCTCTTCGACCATACATATT 1271  
 QY 342 AlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArgAlaArg 361  
 Db 1272 GCTGAAAAGTACGATGATGAGAGAGCTTTATGAGAGAACAAATCCAGAGGCGCAAGA 1331  
 QY 362 LysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpPro 381  
 Db 1332 AAGTCTTCGATGAGGACAGAGAGAGAGTTTACACAACTAGCTCAGCTATGACGATGCT 1391  
 QY 382 ValLeuThrAlaLeuLeuTyrArgLysIleLeuAspGluIleGlnAlaAsnAspTyr 401  
 Db 1392 GTATGCGCATCTTCTGTTGACCGCCAAATATCGACGAGATGGAAGCCATATCATAC 1451  
 QY 402 AsnAsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIle 421  
 Db 1452 AACCACTTCACAAAGAGCTTATGAGCAACCAAGAAAGACTTAATTTCTTACCTATT 1511  
 QY 422 AlaTyrAlaLysSerLeuValProProAsnArgThr-SerSerProLeuAlaLys 439

Db 1512 GCTATGCAAAATCTCTGTGCCCCCTACAGAACTTGTACCTTAGCTAAG 1566  
 RESULT 3  
 US-09-371-307-75  
 / Sequence 75, Application US/09371307A  
 / Patent No. US20020053095A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Brown, Sherri M.  
 / APPLICANT: Heck, Gregory R.  
 / APPLICANT: Piller, Kenneth J.  
 / APPLICANT: Kishore, Ganesh M.  
 / APPLICANT: Ellich, Ted D.  
 / APPLICANT: Logusch, Eugene W.  
 / APPLICANT: Rao, Sudabathula  
 / APPLICANT: Ream, Joel E.  
 / APPLICANT: Logusch, Sherry J.  
 / TITLE OF INVENTION: Methods for controlling gibberellin levels  
 / FILE REFERENCE: MOBT:216  
 / CURRENT APPLICATION NUMBER: US/09/371,307A  
 / CURRENT FILING DATE: 1999-08-10  
 / NUMBER OF SEQ ID NOS: 89  
 / SOFTWARE: PatentIn Ver. 2.0  
 / SEQ ID NO 75  
 / LENGTH: 1239  
 / TYPE: DNA  
 / ORGANISM: Lycopersicon esculentum  
 US-09-371-307-75  
 Alignment Scores:  
 Pred. No.: 7.02e-208 Length: 1239  
 Score: 1737.00 Matches: 342  
 Percent Similarity: 86.37% Conservative: 32  
 Best Local Similarity: 78.98% Mismatches: 35  
 Query Match: 76.52% Indels: 24  
 DB: 9 Gaps: 3  
 US-09-847-081B-2 (1-440) x US-09-371-307-75 (1-1239)  
 QY 3 MetSerValAlaLeuLeuTyrValValSerProThrSerGluValSerAsnGlyThrGly 22  
 Db 1 ATGCTGTGGCTGTTATGAGGTGTTCTCTCT--TCGACGCTCAAAATGGGCAAGT 57  
 QY 23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42  
 Db 58 TTCAATGATATAGTCCGGAAGGAAACCGTTTATTGATTATCG--AGG 105  
 QY 43 AspArgAsnLeuMetTrpAsnGlyArgIleLysLysGlyGlyArgGlnArgTrpAsnPhe 62  
 Db 106 CATAGAAATTTGGTGTCCAAATGAGAGATCAATATGAGGT----- 144  
 QY 63 GlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluLeuGly 82  
 Db 145 -----GGTGAAGAAAGCAAACTAATTAATGGA 168  
 QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThrValSer 102  
 Db 169 CGGAATTTTCTGTCGCTGCTATTTTGGCTACTCATCTGGAGAACGAGCATGACA 228  
 QY 103 SerGluLysValTyrAspValValLeuValGlnAlaAlaLeuValTyrAspGlnLeu 122  
 Db 229 TCGGAACAGAGTGTATGATGTGTTTGAAGGACGAGCGCTGTGTAAGGCACTG 288  
 QY 123 ArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeuGlyLeu 142  
 Db 289 AGATCTACCAATGAGTTTGAAGTGAAGCGGATATTAATTCGCGGGAATTTGGGCTTG 348  
 QY 143 LeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThrPheTyr 162  
 Db 349 TTGACTGAACATATGATAGGTGTGTGATATGTGACAGGTATGCAAGACGTTTAC 408  
 QY 163 LeuGlyThrLysLeuMetThrProGluArgArgAlaIleTyrAlaIleTyrValTrp 182

Db	409	TTAGAACTAATGCTAATAGACTCCGAGAGAAAGAGGCGTATCTGGGCAATATATGTATG	468
Oy	183	CysatgagtgtttaraspgluleuValaspGlyProaenAlaserHisIleThrProGlnAla	202
Db	469	TGCAGAAACAAGATGAACCTTGTTATGCGCCCAACGCATCATATATTACCCGGCAGACC	528
Oy	203	LeuAspAcGTTPGlnThrArgleuGlnAspIlePheSerGlyArgProPheAspMetLeu	222
Db	529	CTAGATTAGGTGGAAATATGGCTTAAGATGTTCTTAATGGGGGGCCATTGACATGGCTC	588
Oy	223	AspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArgasp	242
Db	589	GATGGTGGCTTTGTCCGATACAGTTCTTAACCTTCAGATTGATATTCAGCATTCAGAGAT	648
Oy	243	MetIleGluGlyMetArgMetAspLeuTryptylsSerArgTyrLysThrPheAspGluLeu	262
Db	649	ATGATGTGAAGCAATCCGATAGACTTGGAAATATCAGATACAAATCACTTCACACACTA	708
Oy	263	TyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGly	282
Db	709	TACCTTATATGTTATATATGTTGCTGTACGCTGGTGGTGTGATAGTGTCAATTATGAGT	768
Oy	283	IleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGly	302
Db	769	ATCGCCCTCGAATCAAGGCAACAACAGACGCTATATTAAGCTGCTTGGCTCTGGGG	828
Oy	303	LeuAlaAsnGlnLeuThrAsnIleLeuArgaspValGlyGlnAspAlaArgArgGlyArg	322
Db	829	ATCGCAATTCATTTAATTCACTCACTCAGACGATGTTGGAGAAATGCCAGAAAGAGA	888
Oy	323	ValTyrLeuProGlnAspGlnLeuAlaGlnAlaGlyLeuSerAspGlnAspIlePheAla	342
Db	889	GCTACCTGCTCAAGATGAATTTACACACGCGAGCTTATCCGATGAAGAATATATTGCT	948
Oy	343	GlyArgValThrAspLysTyrArgAsnPheMetLysLysGlnIleGlnArgAlaArgLys	362
Db	949	GGAAGGCTGACCGATTAATGGAGATCTTTATGAAGAAACAAATACATAGGCCAAGAAG	1008
Oy	363	PhePheAspGluSerGlyLysGlyValThrGluLeuAspSerAlaSerArgTribProVal	382
Db	1009	TTCTTTGATGAGGACAGAAATGGCGTACAGAAATTGAGCTCAGTAGTATATTCCCTGTA	1068
Oy	383	LeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGlnIleGlnAlaAsnAspTyrAsn	402
Db	1069	TGGGCATCTTGGCTCTTGTAACCGCAAAATATCATAGATGAGATTGAAGCCAAATCATACAC	1128
Oy	403	AsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIleAla	422
Db	1129	AACCTTCAACAAGACGATATGTGAGCAATCAAGAAGCTTGATTTGCATTACTTATTGCA	1188
Oy	423	TyrAlaLysSerLeuValProProAsnArgThrSerSer 435	
Db	1189	TATGCAAAATCTTGTGCTCTCTTCAAAAACCTGCTCT 1227	
RESULT 4			
US-10-401-321-75			
: Sequence 75, Application US/10401321			
: Publication No. US20030233679A1			
: GENERAL INFORMATION:			
: APPLICANT: Brown, Sherri M.			
: APPLICANT: Heck, Gregory R.			
: APPLICANT: Pillier, Kenneth J.			
: APPLICANT: Kilshore, Ganesh M.			
: APPLICANT: Ellich, Ted D.			
: APPLICANT: Logusch, Eugene W.			
: APPLICANT: Rao, Sudabathula			
: APPLICANT: Ream, Joel E.			
: APPLICANT: Logusch, Sherry J.			
: APPLICANT: Baerson, Scott R.			
: TITLE OF INVENTION: Methods for Controlling Gibberellin Levels			
: FILE REFERENCE: 1189, 0216, DIVISO1 (MOBRT-216--1)			
: CURRENT APPLICATION NUMBER: US/10/401,321			
: CURRENT FILING DATE: 2003-03-27			

NUMBER OF SEQ. ID NOS: 89			
SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 75			
; LENGTH: 1239			
; TYPE: DNA			
; ORGANISM: Lycopersicon esculentum			
US-10-401-321-75			
Alignment Scores:			
Pred. No.:	7.02e-208	Length:	1239
Score:	1737.00	Matches:	342
Percent Similarity:	86.37%	Conservative:	32
Best local Similarity:	78.98%	Mismatches:	35
Query Match:	76.52%	Gaps:	24
DB:	12	Indels:	3
US-09-847-081B-2 (1-440) x US-10-401-321-75 (1-1239)			
QY	3 MetSerValAlaIeuIeuTrpValValSerProThrSerGluValSerAsnGlyThrGly	22	
DB	1 ATGTCTGTGGCTTGTATGGGTGTGTTCTCT--TGTGACGTCTCAAAATGGGCAAGT	57	
QY	23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg	42	
DB	58 TTGATGGATCAAGTCCGGAGGGAACCGTTTTTTTGATTCATCG-----AGG	109	
QY	43 AspArgAsnIeuMetTrpAsnGlyArgIleIysIysGlyIaArgIlnaArgTrpAsnPhe	62	
DB	106 CATAGCAATTTGGTGTCCAATGAGAGAAATCAATAGAGGT-----	144	
QY	63 GlySerLeuIleAlaAspProArgTrpIserCysIeuGlyIysSerArgThrGluIleGly	82	
DB	145 -----GGTGGAAAGCAAACTAATATATGGA	168	
QY	83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyIleMetThrValSer	102	
DB	169 CGGAATTTTCTGTACGGCTCTCTATTTTGGCTACTCATCTGGAGAACGACATGACA	228	
QY	103 SerGluIlyAlaValTyrAspValValIeuIySglnAlaAlaIeuValIysArgGlnIleu	122	
DB	229 TCGGAACAGATGGCTATGATGTGGTTTGAAGCGACGACCTTGTTGAAGAGCACTGG	288	
QY	123 ArgSerThrAspAspIeuGluValIysProAspIleValProGlyAsnIleuGlyIleu	142	
DB	289 AGATCTTCAACATGATTAAGTTAAGCTGAAGCCGATATACCTATTTCCGGGGAATTTGGGCTTG	348	
QY	143 LeuSerGluAlaIleTyrAspArgCysGlyGlnValCysAlaGlyIlyrAlaIleTyrPheTyr	162	
DB	349 TTGATGGAAACATATATATAGTGTGTGTGTAAGTATGTGACAGTATGCMAAGACCTTTAAC	408	
QY	163 LeuGlyThrIlySleuMetThrProGluArgArgAlaIleIleTyrAlaIleTyrValTyrP	182	
DB	409 TTGAGAACATATGCTATATGACTCCCGAGAGAAAGGCTTACTGGCAATATATATGTATGG	468	
QY	183 CysArgArgThrAspGluIeuValAspGlyProAlaAlaSerHisIleThrProGlnAla	202	
DB	469 TGCAGAGAAACAGATTAACCTTTGATGCGCCCAACGATCATATATTTAACCCCGGACGCC	528	
QY	203 LeuAspArgTrpGluIleThrArgIleuIlyAspIlePheSerGlyArgProPheAspMetIleu	222	
DB	529 CTGATATAGTGGGAAATATAGGCTTAAAGATGTGTTTCAATGGGCGCCATTTTGACATGCTC	588	
QY	223 AspAlaAlaIeuSerAspThrValSerArgPheProValAspIleGlnProPheArgAsp	242	
DB	589 GATGGTGCTTTGTCCGATACAGTTTCTAACTTCCAGTTGATATTCAGCCATTCCAGGAT	648	
QY	243 MetIleGluGlyMetArgMetAspLeuTrpIlySerArgIlyIlySerPheAspGluIleu	262	
DB	649 ATGATTTGAAGAGATCGGTATGAGCTTGAAGAAATCGAATATCAAAAACTTCAGAGAACTA	708	
QY	263 TyrIleuTrpCysIleTyrTrpValAlaGlyThrValGlyIleuMetSerValProValMetGly	282	
DB	709 TACCTTATATGTATATATATGTCTGTACGTTGGTGTGAAGAGTTCCAATATATGAGGT	768	

QY	283	lelaaPProGiuserlyValaIaThrThcIuserValIyThaenIaaIaleuAlaIeugIy	302
Db	769	ATGCGCCCTGATCAAAAGCAACAAGAGCGGATATATATATGCTTGGCTTGGGG	828
QY	303	LeuAlaengInLeuThaSnIleuArygaSpValGIyGIuAspAlaArArGIYArG	322
Db	829	ATCGCAATTCATTAATCTTACCTACTCTACAGATGTGGAGAAGATCCGAAGAAGAA	888
QY	323	ValIyTLeuPProGInaSPGIuLeuAlaGInIaGIyLeuSerAaPGIuAsPIIePheAla	342
Db	889	GTCTACTTGGCCCTCAAGATGAAATTAGCAAGAGCAGGTCTATCCGATGAAGATATATTGGCT	948
QY	343	GIyArGValIThAAsPIySThArAAsnHemelysIySGInIeGInArGAlaArGIyS	362
Db	949	GGAGGGTGACCGATTAATGGAGAACTTTTATGAAGAACAAATATCATAGGCAAGAAAG	1008
QY	363	PhePheAspGIuSerGIuLySGIyValIThrGIuLeuAspSerAlaSerArGTIPProVal	382
Db	1009	TTCTTTGATGAGGCAAGATGGCGGTGACAGATTAAGACTCAGCTAGTATATTCCTGTATA	1068
QY	383	LeuThraIaLeuLeuLeuTyThArgLySIleLeuAspGIuIeGIuIaAsnAsPTYrAsn	402
Db	1069	TGGGCATTTTGGCTGTGTACCGCAAAATACTAGATGAGATTGAAGCCAAAGACTTCAAC	1128
QY	403	AsnPheThrArGArGAlaIyValSerIySProLySIyLeuLeuThLeuProIleAla	422
Db	1129	AACTTCACAAAGAGACATATGTAGCAAAATCAAAAGAAAGTTGATTCCTTACTTATGCA	1188
QY	423	TyralAlaYSerLeuValProProAsnArGThrSerSer	435
Db	1189	TATGCAAAATCTATGTGCTCTTCACAAAATCTGCCTCT 1227	

## RESULT 5

```

US-09-938-842A-729
; Sequence 729, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 729
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-729

```

**Alignment Scores:**

Pred. No.:	2,376,183	Length:	1269
Score:	1542.00	Matches:	318
Percent Similarity:	81.86%	Conservative:	43
Best Local Similarity:	72.11%	Mismatches:	51
Query Match:	67.93%	Indels:	30
DB:	10	Gaps:	8

US-09-847-081B-2 (1-440) x US-09-938-842A-729 (1-1269)

```

Oy      1 MetSerMetSerValAlaLeuLeuTyrValVal-----SerProThrSerGluValSer 18
        |||||  |||||:::|||||
Db      1 ATGCTCTTCTTCGTAGCAGTGTATGGTGTCTACTTCTTCTCMAATCCAGACCAATG 60

```

OY		19	AseNglYthrGlyLeuLeuAaspSerValArgGlueLysAsnArgValPheValSerSerArg	38
Dd		61	AACATTTGGTGGTGGTA-----AGGGTCTTAGAAATCTTCCTAGA	99
OY		39	PheLeuAlaArgAspArgAsnLeuMetTrpAsnGlyArgIleLysLysGly-GlyArgGI	58
Dd		100	CTGTTCCTCCTTGCCAG-----AATCAGAGACTCAACAAGCTAAGAAGAG	147
OY		58	nArgTrpAsnDheGlySerLeuIleAlaAspProArgTySerCysLeuGlyGlySerAr	78
Dd		148	CAGATACCACCTTGAGAGTCTTCCTT-----TTGTATAGGAACCGA	186
OY		78	gThrGluLysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyI	98
Dd		187	AGTtGAAAGAATT-GGTGTT-----GGTCTTCAAGCTTATACAGACTCTTCCTGAGA	239
OY		98	uMetThrValSerSerglulysLysValTyAspValValLeuLysglnAlaAlaLeuVa	118
Dd		240	GATAGCTCTTCATCTGAGAGAAGAGGTTTACAATGTTGTGTGAACAAGCTGGTTGGT	299
OY		118	LlysArgGlnLeuArgSerThrAsp--AspleuGluValLys-----ProAspILeVa	135
Dd		300	GAACAACAAGTAAAGGCTCTCTCTTATGACCTTGATGTGAAGAACCAACAAGATGTTGT	359
OY		135	IValProGlyAsnDheGlyLeuLeuSerglulAlaTyAspArgCysGlyGluValCysAl	155
Dd		360	TCTTCCTGGAGAGTTGAGTTGAGTTGGTGGTGAAGCTTATGATGATGCAGCGGTGAAGTTGGCG	419
OY		155	aGluTyAlaLysThrPheTyLeuGlyThrLysLeuMetThrProGluArgArgArgAl	175
Dd		420	TGAATATAGCTAAAGCTTTTATCTTGGAACCTTGCTTATGACACCCGAAGGCCAAAGGC	479
OY		175	AlaIleTrpAlaIleTyArgValTrpCysArgArgThrAspGluLeuValAspGlyProAsnAl	195
Dd		480	GATTTGGCAATCTCAAGTTGGGTGATGAGAACTGAAGAATTGTGAGATGGGCCAAATGC	539
OY		195	aSerHisIleThrProGlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSe	215
Dd		540	TTCACTATTAATCTCCCATGGCTTTAGATAATGGGAAGCAAAGTTAGAAAGATCTTTCCG	599
OY		215	rGlyArgProPheAspMetLeuAspAlaAlaLeuSerAspThrValSerArgPheProVa	235
Dd		600	TGGTCGTCCTTCATATGCTTGATGCTGCTCTCGCATATACATTCCTGATCCCGGT	659
OY		235	IAspLIleGlnProPheArgAspMetIleGluGlnMetArgSerAspLeuTrpLysSerAr	255
Dd		660	CGATATTCAGCAATTTGAGACATGATCGAAGGATGAGATGAGACTTTGAAGAAATCGAG	719
OY		255	gTyLysThrPheAspGluLeuTyLeuTyCysTyTyValAlaGlyThrValGlyLe	275
Dd		720	ATACACAGAACTTCGATATCTATCTACCTTACTGCTACTACGTCCTGGAACCGTCGATT	779
OY		275	uMetSerValProAlaMetGlyIleAlaProGluSerLysAlaThrThrGluSerValTy	295
Dd		780	GATAGCGCTTCGGATTATGGGAATCGATCTCAATGTCAAAGCAACAACCGAAATGTTTTA	839
OY		295	rAsnAlaAlaLeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGI	315
Dd		840	CAAGCGTCGCTTGCGCCCTTGSTATAGCCAATCAGCTTACTAATCAGAGAGCGTAGG	899
OY		315	yGluAspAlaArgArgGlyArgValTyLeuProGlnAspGluLeuAlaGlnAlaGlyLe	335
Dd		900	CGAAGAATGCCGAAGAAGAGGTTTATCTGCCACAGATGAATTTGCTCAGCGTGGCTCT	959
OY		335	uSerAspGluAspLIlePheAlaGlyArgValAlaThrAspLysTrpArgAsnPheMetLysLy	355
Dd		960	TTTCAGATGAACATATTCGCCGGAAGAAAGTACTGATTAATGAGAAATCTCATGAHAAT	1019
OY		355	sGlnIleGlnArgAlaArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAs	375
Dd		1020	GCACCTTAAACGAGCAAGAAATGTTCTTCGACGAGAGCTGAAGAAAGGCGCTCACCCGAGCTCAG	1079

Db 1020 GCAGCTTAAACGACGACGATGTTCTTCGACGAGCTGAGAGAGGCGTCACCGAGCTAG 1079

```

Qy 375 pserAlaSerArgTTPProValLeuThraAlaLeuLeuTyrArgLysIleLeuAspG1 395
Db 1080 TGCCCTAGACAGATGGCGTGTATGGCGTTCATTTGCTATTGTACAGAGAACTACTGACGA 1139
Qy 335 uIlleGluAlaasnAapTYrAsnAspPheThraArgAlaIlyrValSerLysProLysI 415
Db 1140 GATTAAAGCGAATGATTCACAAATTTTAACTAAGAGAGCTTATGTGGGAAAGTCMAA 1199
Qy 415 sLeuLeuThraLeuProIleAlaIlyrAlaLysSerLeuValProProAsnArgThrSerSe 435
Db 1200 AATTGCAGCTTTCCTTGGCTTATGCTTAATCACTACTA-----AAGACTTCAAG 1250
Qy 435 r 435
Db 1251 T 1251

RESULT 6
US-10-149-759-53
; Sequence 53, Application US/10149759
; Publication No. US20030157592A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reinold, Andreas
; APPLICANT: Cirus, Petra
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins
; TITLE OF INVENTION: involved in the synthesis of tocopherols and
; FILE REFERENCE: BASF/NAE 1333/99 PCT/US
; CURRENT FILING DATE: 2002-10-17
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: WordPerfect version 6.1
; SEQ ID NO 53
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Physcomitrella patens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(397)
; OTHER INFORMATION: 02_ppprocl_046_a07rev
US-10-149-759-53

Alignment Scores:
Pred. No.: 9,34e-49 Length: 684
Score: 468.00 Matches: 93
Percent Similarity: 82.31% Conservative: 14
Best Local Similarity: 71.54% Mismatches: 23
Query Match: 20.62% Indels: 0
Gaps: 0

US-09-847-081b-2 (1-440) x US-10-149-759-53 (1-684)
Qy 310 IleLeuArgAspValGlyGluAspAlaArgArgGlyArgValTyrLeuProGlnAspGlu 329
Db 5 ATCCGAGGAGATGTTGAAGAAGATGACGCCGTGGAGAGTATACCTCCACAGGATGAA 64
Qy 330 LeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAlaGlyArgValThrAspLysTrp 349
Db 65 CTGGCAGCTTTCGGTCTGTCTGATGACGACATTTTTCGCGAAAGTTACTGATTAAGG 124
Qy 350 ArgAsnPheMetLysLysGlnIleGlnArgAlaArgLysPhePheAspGluSerGluLys 369
Db 125 AGGGCATTCATGAAAGACCAATTAAGAGCTAGAGCTTCTTGTGGAGGCTGAGAA 184
Qy 370 GlyValThrGluLeuAspSerAlaSerArgTTPProValLeuThraAlaLeuLeuTyr 389
Db 185 GGTGTACGTGACGACAAAGACAGCTGCTGCTGTGTGTGCTGCGCCCTCATTTCTTAC 244
Qy 390 ArgLysIleLeuAspGluIleGluAlaasnAapTYrAsnAspPheThraArgAlaTyr 409

```

```

Db 245 CAGCAAAATTCGAGCGCCATTGAAAGCCAGATTACATTAATTCAACAAAAGAGCTTAC 304
Qy 410 ValSerLysProLysLysLeuLeuThraProIleAlaIlyrAlaLysSerLeuValPro 429
Db 305 GTAGCAAGAGTGAGAAAGAGCTGCTTCTTACTTATGCTTATGACAGAGCTTGTTCA 364
Qy 430 ProAsnArgThrSerSerProLeuAlaLys 439
Db 365 CTTCAGATGACCTTCCAGGTTAGACGT 394

RESULT 7
US-10-156-761-1013
; Sequence 1013, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHITYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT FILING DATE: 2002-05-29
; PRIOR FILING DATE: 2001-05-30
; PRIOR FILING DATE: 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1013
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1026)
US-10-156-761-1013

Alignment Scores:
Pred. No.: 2,04e-36 Length: 1026
Score: 372.00 Matches: 101
Percent Similarity: 51.75% Conservative: 47
Best Local Similarity: 35.31% Mismatches: 116
Query Match: 16.39% Indels: 22
Gaps: 7

US-09-847-081b-2 (1-440) x US-10-156-761-1013 (1-1026)
Qy 143 LeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThrPheTyr 162
Db 46 CTGGCAGAGGCTTACACGACCTGCGCGGCTCAACCGCGACCGGCAAGACCTACTTC 105
Qy 163 LeuGlyThrLysLeuMetThrProGluArgArgAlaIleThrAlaIleTyrValTyr 182
Db 106 CTGCCACACAGATGCTGCGCGGTGAGAGCGCGCGCGCGCTGACGCGCTTACGARTTC 165
Qy 183 CysArgArgThrAspGluLeuValAsp-----GlyProAsnAlaSer 196
Db 166 GCACGCTGGCGCGACGATGTCGACTCCCTGGACACCAACCGTGGCCCC----- 216
Qy 197 HisIleThrProGlnAlaLeuAspArgTTPGluThrArgLeuGlu---AspIlePheSer 215
Db 217 GCTGTGCTCTCGCGCGGCTGCGCGCTCCAGAGAGCTTGGCGAGCGGAGCTGAG 276
Qy 216 GlyArgProPheAspMetLeuAspAlaAlaLeuSerAspThrValSerArgPheProVal 235
Db 277 GGAACACGACCGAGCGGTGCTGCTGCGCTGCGACAGACGCGCGCGGTACGCATC 336
Qy 236 AspIleGlnProPheArgAspMetIleGluGlyMetArgMetAspLeuTyrLysSerArg 255

```



```

Db      337 GACACACAGACTTCATGATCTTATGAGCGCCATCGACGACGACCTGAGGTACCGGC 396
Qy      256 TTYLSTHrPheAspGluLeuTYrLeuTYrCysTYrTYrValAlaGlyThrValGlyLeu 275
      397 TACGAGACTTACCGACCTGCGCGGCTTACATGCACAGCTTCCGCGCGGATCGGGCTG 456
Qy      276 MetSerValProValMetGly---IleAlaProGluSerIleValAlaThrThrGluSerVal 294
      457 CAGATCTCTCGGTCTCGGAGCCGTGTCTCCGCGTGAAGAGCGCCACCCAC----- 510
Qy      295 TYRAsnAlaAlaLeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspVal 314
      511 -----GCGCGGCGCTCGGAGTGGCGGTTCACCTGACCAACTTCTCGCGGACGCG 561
Qy      315 GtGtLAspAlaArgArgGlyArgValTYrLeuProGlnAspGluLeuAlaGlnAlaGly 334
      562 GGGAGAGACTTGACCGCGCGCGCTCTATCTGCGCGGACCTGTGGCGCGCCACGCGC 621
Qy      335 LeuSerAspGluAspIlePhe-----AlaGlyArgValThrAspIleSTrp 349
      622 GTC---GACCGGAGACTGCTGCGGCTGAGAGAGGACACCGCGCGCGGACCGCGTATC 678
Qy      350 ArgAsnPhemeIleTyseGlnIleGlnArgAlaArgIlePheAspGluSerGluIys 369
      679 ACCCGTGCCTCAAGGCCGCTTGAAGGCGCTGACCCGCGGTGTGTACCGCAGAGGCGCGCG 738
Qy      370 GlyValThrGluLeuAspSerAlaSerArgTrpProValLeuThrAlaLeuLeuLeuTYr 389
      739 GGGCTTGCATGCTGACCCGCGGCGCGCGCGCGCGTGCATCCGAGCGCGTTCGTGTC 798
Qy      390 ArgIleValLeuAspGluIleGlnAlaAsnAspTYrAsnAsnPhenThrArgAlaTYr 409
      799 GCGGAGATCTTGAGCCCATCGCGCAGACGCGGTACGCGGTGTCCACCGCGTCCGCG 858
Qy      410 ValSerIleProIleIys 415
      859 GTGCGCGCGCGCGCGCT 876
Db

```

RESULT 8  
US-10-156-761-1  
; Sequence 1, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OKURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; PRIOR FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 1  
; LENGTH: 9025608  
; TYPE: DNA  
; ORGANISM: Streptomyces avermitilis  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (4187715)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
US-10-156-761-1

Alignment Scores:  
Pred. No.: 2,596-30 Length: 9025608  
Score: 372.00 Matches: 101  
Percent Similarity: 51.75% Conservative: 47

```

      162 1291904
Qy      143 LeuSerGluAlaTYrAspArgCysGlyValAlaCysAlaGluTYrAlaIleThrPheTYr 162
      1291845 CTGCCAAGAGCTTACAGCACTGCGCGCGGCTCAACGCGCGGACAGGACGATCTTC 1291904
Db      163 LeuGlyThrIleLeuMetThrProGluArgArgAlaIleTPrAlaIleTYrValTPr 182
      1291905 CTGCGCACACAGACTGCTCGCGGTCTGACGCGCGCGCGCGCTGACGCGCTTACCGAATTC 1291964
Qy      183 CysArgGlyThrAspGluLeuValAsp-----GlyProAsnAlaSer 196
      1291965 GCAGCTGGCGCGGACGACATGCTGATCTCCCTGACACACCGCTGCGCC----- 1292015
Qy      197 HisIleThrProGlnAlaLeuAspArgTrpGluThrArgLeuGlu---AspIlePheSer 215
      1292016 GCTGTGCGCTGCGCGCGCGCTGCGCGCTCCACAGAGAGCTTCCGAGCGGACTGCTGAG 1292075
Qy      216 GlyArgProPheAspMetLeuAspAlaAlaLeuSerAspThrValSerArgPheProVal 235
      1292076 GGACACACACACCGACCGCGGTGCTCTGCGCTGCGCGGACGCGCGCGGTACCGCATTC 1292135
Qy      226 AspIleGlnProPheArgAspMetIleGlnIleMetArgMetAspLeuTrpIleSerArg 255
      1292136 GACCAACAGCACTTATGACTTCTCATGACGCCATGACGCGACGACGCTGAGGTACCGGAC 1292195
Qy      256 TYLSTHrPheAspGluLeuTYrLeuTYrCysTYrTYrValAlaGlyThrValGlyLeu 275
      1292196 TACGAGACTTACCGGACCTGCGCGGCTTACATGACAGGTTGCGCGCGGTATCGGCTG 1292255
Qy      276 MetSerValProValMetGly---IleAlaProGluSerIleValAlaThrThrGluSerVal 294
      1292256 CAGATCTGCGGTGCTCGGAGCCGTGTCTCCGCGTGAAGAGCGCGCACCCAC----- 1292309
Qy      295 TYRAsnAlaAlaLeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspVal 314
      1292310 -----GCGCGGCGCTCGGAGTGGCGGTTCACGCTGACCAACTTCTCGGAGACGTCG 1292360
Qy      315 GlyIleAspAlaArgArgGlyArgValTYrLeuProGlnAspGluLeuAlaGlnAlaGly 334
      1292361 GCGGAGAGCTTGAGCCGCGCGCGCGGTCTATCTGCGCGGACCTTTCGCGCCACGCGC 1292420
Qy      335 LeuSerAspGluAspIlePhe-----AlaGlyArgValThrAspIleSTrp 349
      1292421 GTC---GACCGGAGACTCTGCGGCTGAGAGAGGACACCGCGCGCGGACCGCGTATC 1292477
Qy      350 ArgAsnPhemeIleTyseGlnIleGlnArgAlaArgIlePheAspGluSerGluIys 369
      1292478 ACCCGTGCCTCAAGGCCGCTTGAAGGCGCTGACCCGCGGTGTGTACCGGAGAGCGCGCG 1292537
Qy      370 GlyValThrGluLeuAspSerAlaSerArgTrpProValLeuThrAlaLeuLeuLeuTYr 389
      1292538 GGGCTTGCATGCTGACCCGCGGCGCGCGCGCGCGTGCATCCGAGCGCGTTCGTGTC 1292597
Qy      390 ArgIleValLeuAspGluIleGlnAlaAsnAspTYrAsnAsnPhenThrArgAlaTYr 409
      1292598 GCGGAGATCTTGAGCGCGGACGCGGTACGCGGTGTCCACCGCGTCCGCG 1292657
Qy      410 ValSerIleProIleIys 415
      1292658 GTGCGCGCGCGCGCGCT 1292675
Db

```

Beet Local Similarity: 35.31% Mismatches: 116  
Query Match: 16.39% Indels: 22  
DB: 15 Gaps: 7

US-09-847-081B-2 (1-440) x US-10-156-761-1 (1-9025608)

RESULT 9  
US-10-156-761-1642  
; Sequence 1642, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OKURA, SATOSHI  
; APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN  
 APPLICANT: HORIKAWA, HIROSHI  
 APPLICANT: SHIBA, TADAYOSHI  
 APPLICANT: SAKAKI, YOSHIYUKI  
 APPLICANT: HATTORI, MASAHIRA  
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 FILE REFERENCE: 249-262  
 CURRENT APPLICATION NUMBER: US/10/156,761  
 CURRENT FILING DATE: 2002-05-29  
 PRIOR APPLICATION NUMBER: JP 2001-204089  
 PRIOR FILING DATE: 2001-05-30  
 PRIOR APPLICATION NUMBER: JP 2001-272697  
 PRIOR FILING DATE: 2001-08-02  
 NUMBER OF SEQ ID NOS: 15109  
 SEQ ID NO 1642  
 LENGTH: 948  
 TYPE: DNA  
 ORGANISM: Streptomyces avermitilis  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (1)..(948)  
 US-10-156-761-1642

Alignment Scores:  
 Pred. No.: 5,766-36 Length: 948  
 Score: 368.00 Matches: 96  
 Percent Similarity: 52.45% Conservative: 43  
 Best Local Similarity: 36.23% Mismatches: 104  
 Query Match: 16.21% Indels: 22  
 Gaps: 6

US-09-847-081b-2 (1-440) x US-10-156-761-1642 (1-948)

Oy 146 AlAtYrAspArgCyGgLyGluValCyAlaGluTyrAlaYsThrPheTyrLeuGlyThr 165  
 Db 52 GCATACAGCTACTGCGAGACCGTCAACCGGCGAGCGCCGCACTTCCCTACGCGCATC 111  
 Oy 166 LysLeuMetThrProGluArgArgAlaIleTTPAlaIleTyrValTTPCySarArg 185  
 Db 112 AGCGTGTGCGCGACGCCCAAGCGCGCGCGATGCGCGCTTACGCTTCCGCGCGCC 171  
 Oy 186 ThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGlnAlaLeuAspArg 205  
 Db 172 GTCCGACGATCGCGAGCGCGCGCTCCGCGCGCGAGCGCGAGCGCGCTCGAGGAC 231  
 Oy 206 TrpGluThrArgLeuGluAspIlePheSerGlyArg-----ProPheAspMet 221  
 Db 232 ACCCGGGCGCTGTCCCGGTGCGTGAACGGCGGTGAGACGAGACGACGACGCC 291  
 Oy 222 LeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArg 241  
 Db 292 GTCCGCGTCCCTTGGCGACGCCCGCGAGCGAGCTTCCGATCCCGCTCGCGCGCTCAGC 351  
 Oy 242 AspMetIleGluGlyMetArgMetAspLeuTyrPlySerArgTyrLeuThrPheAspGlu 261  
 Db 352 GAATGATCGACGCGGTCTCATGACGCGCGCGCGAGACCTTACGAGACCTGGAGACGAC 411  
 Oy 262 LeuTyrLeuTyrCyGtyTyrValAlaGlyThrValGlyLeuMetSerValProValMet 281  
 Db 412 CTCAGAGTCTACTGCGCTGTGTGCGACGGCGCATCGCGCGCTGCTCGCGCTGTTC 471  
 Oy 282 GlyIleAlaProGluSerValaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeu 301  
 Db 472 GGCACGGAACCGGGCG 528  
 Oy 302 GlyLeuAlaAsnGluLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGly 321  
 Db 529 GACCTCGCTCTCCACTCAGCAACATCTCGCGCGAGCGAGAGCGCGCGCGCGCGCGCG 588  
 Oy 322 ArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSer-----AspGlu 338  
 Db 589 GCGACATCTATGCGCGCGCGAGCAGCTCGCAAAATTGCGCTGCTCGCGCGGTTCAGCGG 648

Oy 339 AspIle-----PheAlaGlyArgValThrAspLysTTPArgAspMet 352  
 Db 649 CCAGTCCACCGAGGAGATCGCACTTCCGCGCGCTGCTG-----CACTTC 693  
 Oy 353 MetLysIleGlnIleGlnAlaArgLysPheAspGluSerGlyValThr 372  
 Db 694 -----GAGGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 744  
 Oy 373 GluLeuAspSerAlaSerArgTTPProValLeuThrAlaLeuLeuLeuTyrArgLysIle 392  
 Db 745 ATGCTGACCG 804  
 Oy 393 LeuAspGluIleGlu 397  
 Db 805 CTCGATCGCATCGAG 819

RESULT 10  
 US-10-156-761-1/c  
 Sequence 1, Application US/10156761  
 Publication No. US20030119018A1  
 GENERAL INFORMATION:

APPLICANT: OMIURA, SATOSHI  
 APPLICANT: IKEDA, HARUO  
 APPLICANT: ISHIKAWA, JUN  
 APPLICANT: HORIKAWA, HIROSHI  
 APPLICANT: SHIBA, TADAYOSHI  
 APPLICANT: SAKAKI, YOSHIYUKI  
 APPLICANT: HATTORI, MASAHIRA  
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 FILE REFERENCE: 249-262  
 CURRENT APPLICATION NUMBER: US/10/156,761  
 CURRENT FILING DATE: 2002-05-29  
 PRIOR APPLICATION NUMBER: JP 2001-204089  
 PRIOR FILING DATE: 2001-05-30  
 PRIOR APPLICATION NUMBER: JP 2001-272697  
 PRIOR FILING DATE: 2001-08-02  
 NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 1  
 LENGTH: 9025608  
 TYPE: DNA  
 ORGANISM: Streptomyces avermitilis  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (4187715)  
 OTHER INFORMATION: a, t, c, g, other or unknown  
 US-10-156-761-1

Alignment Scores:  
 Pred. No.: 8,246-30 Length: 9025608  
 Score: 368.00 Matches: 96  
 Percent Similarity: 52.45% Conservative: 43  
 Best Local Similarity: 36.23% Mismatches: 104  
 Query Match: 16.21% Indels: 22  
 Gaps: 6

US-09-847-081b-2 (1-440) x US-10-156-761-1 (1-9025608)

Oy 146 AlAtYrAspArgCyGgLyGluValCyAlaGluTyrAlaYsThrPheTyrLeuGlyThr 165  
 Db 2025896 GCATACAGCTACTGCGAGACCGTCAACCGGCGAGCGCCGCACTTCCGCTACGCGCATC 2025837  
 Oy 166 LysLeuMetThrProGluArgArgAlaIleTTPAlaIleTyrValTTPCySarArg 185  
 Db 2025836 AGCGTGTGCGCGACGCCCAAGCGCGCGCGATGCGCGCTTACGCTTCCGCGCGCG 2025777  
 Oy 186 ThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGlnAlaLeuAspArg 205  
 Db 2025776 GTCCGACGATCGCGAGCGCGCGCTCGCGCGCGAGCGCGAGCGCGCTCGAGGAC 2025717  
 Oy 206 TrpGluThrArgLeuGluAspIlePheSerGlyArg-----ProPheAspMet 221  
 Db 2025716 ACCCGGGCGCTGTCCCGGTGCGTGAACGGCGGTGAGACGAGACGACGACGCC 2025657

QY 222 LeuApSPAlAalAleuSerAAspThrValSerArgPheProValAspLLeGlnProPheArg 241  
Db 2025656 GTCCGGCGTGCCTTCGGCCAGCGGCCGACGATGTCCCGATCCGCTGCGCGCCTCGAC 2025597

QY 242 AspMetLleGIuGlIyMeArGMeMetAspLeuTrpLysSerArgTYrLYeThrPheaSpGLu 261  
Db 2025596 GAACGTATCGAACGGCGCTCTCATGAGCACTGGCGCGCGACACTACGAACTTGGAACAAC 2025537

QY 262 LeuTYrLeuTYrCYsTYrTYrVALAlaGLYThrVALGlyLeuMetSerValProVALmet 281  
Db 2025536 CTCAAGCTCTACTGCGCGCTGTGTGGCAAGGGCCATCGGCCCGCTTCGCTCGCGCTGTT 2025477

QY 282 GlyLLeAlaProGluSerLyeALaThrThrGluSerVALTYrASnAlAlaleuAlaleu 301  
Db 2025476 GGCACGGAACCGGGCGCCCGCGGC--GCCAGCGCGCGTACAGATACCGCACCTCT 2025420

QY 302 GlyLeuAlaAsnGlnLeuThrASnLlLeuArgAPValGlyLyuASPAlAArgArgGLY 321  
Db 2025419 GGCTCGCTCTCCAGCTACCACTACTCTCCGACATACCGAGAGACCGGAGGGCGGC 2025360

QY 322 ArgVALTYrLeuProGlnAspGLuLeuAlaGlnAlaGlyLeuSer-----AspGLu 338  
Db 2025359 CGCACCTATCTGCGCGCCGACGACCTCGCCAATTGCGCTGCTGCGCGCGGTTCGACCG 2025300

QY 339 Asplle-----PheAlaGlyArgVALThrAspLysTPArgAsnPhe 352  
Db 2025299 CCGATCTCCACCGAGGATCCGACTCTGGGGGCTCTG-----CACTTC 2025255

QY 353 MetLysLysGlnLleGlnArgAlaArgLysPhePheAspGluSerGLuYSGLYValThr 372  
Db 2025254 -----GAAATCCGTCCGGCCCGCCCTTTTCGCGAGGGCTACCGGCTGCTGCC 2025204

QY 373 GluLeuAspSerAlaSerArgTRIProVALLeuThrAlaLeuLeuTYrArgLyslle 392  
Db 2025203 ATGCTGGACCGCGCGAGCGGTGCTGTGCGCGGATGCGGGGATTAACCGCGGCTC 2025144

QY 393 LeuAspGLuLleGln 397  
Db 2025143 CTCGATCGCATCGAG 2025129

RESULT 11  
US-10-128-713A-15  
; Sequence 15, Application US/10128713A  
; Publication No. US20030170847A1  
; GENERAL INFORMATION:  
; APPLICANT: Bramucci, Michael G  
; TITLE OF INVENTION: Genes Involved in Isoprenoid Compound Production  
; FILE REFERENCE: CI-1788  
; CURRENT APPLICATION NUMBER: US/10/128,713A  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 15  
; LENGTH: 945  
; TYPE: DNA  
; ORGANISM: Rhodococcus erythropolis  
US-10-128-713A-15

Alignment Scores:

Pred. No.:	6 71e-35	Length:	945
Score:	359.50	Matches:	99
Percent Similarity:	48.66%	Conservative:	46
Best Local Similarity:	33.22%	Mismatches:	116
Query Match:	15.84%	Indels:	37
DB:	13	Gaps:	8

US-09-847-081B-2 (1-440) x US-10-128-713A-15 (1-945)

QY 140 LenglyLeuLeuSerGluAlaTYrASpArgCYsGLyGluVALCySAlaGluTYrAlalyS 159  
::: ||||| :::: |||||  
Db 1 ATGAACGATTTGTCTGGCTCTCATATAATTCTGCAAGACATGACAGGAAACACAGCGCGA 60

QY 160 ThrPheTYrLeuGlyThrLysLeuMetThrProGluArgArgAlaLleTPAlaLle 179

Db 61 AGGACTTTCTGGCCACTGGGTGCTGCCGAGCCTCGACGCGCGAGTTCCAGCTCTC 120  
Qy 180 TyrValTyrCysArgThrAspGluLeuValAspGlyProAsnIAspSerIle----- 197  
Db 121 TACGATTTGCTCGCGCTCGACGACGCTGTCGACGACACCTCGGATCCCATGAAGA 180  
Qy 198 -----IleThrProGlnIleAspArgTyrGluThr 208  
Db 181 GGCACGCTGCTCGCCGACGCTCGAACGTCGAGCCGTCACCCGACCTCGAAC-CCCACT 237  
Qy 209 ArgLeuGluAspIlePheSerGlyArgProPheAspMetLeuAspIle-----AlaLeu 226  
Db 238 GCGACAGTGAGCTTCCCGCTCGACGATTCCTCTGACTTGAACGCGTACTCTCTGCTTC 297  
Qy 227 SerAspThrValSerArgPheProValAspIleGlnProPheArgAspMetIleGluGly 246  
Db 298 GCGGATGCTGTGAAGACGTTCCGACATTCGCGGTCGACTTCGACGCGCTTTCGAGTCC 357  
Qy 247 MetArgMetAspLeu-----TyrIleSerArgTyrIleThrPheAspGlu 261  
Db 358 ATGCGGATGAGCGCCCGCCGACACCGGAGTGTTCGACCCGCTCTCAACACGATGAGCAG 417  
Qy 262 LeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMet 281  
Db 418 CTTCGCGAGTACATGATGACGCTCCGCGCTGCTGATCGTGTTCGAGATGCTCCGATTCCTC 477  
Qy 282 GlyIleAla---ProGluSerLysAlaThrThrGluSerValTyrAsnAlaIleAla 300  
Db 478 GAGAGTGAAGCTTCCGACGACGAGAGCTGTAAGT-----CCGCGCTCGAAT 522  
Qy 301 LeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArg 320  
Db 523 CTTCGAGAGCGCTTTCAGCTGACCACTTCATCCGCGACGTCGAGTGAAGACCTCGACCG 582  
Qy 321 GlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIle 340  
Db 583 GGACCTGTATCTCCCGCGGCGGAGTTCGCGCATTCGCGGTCGACATCGAGATGCTC 642  
Qy 341 PheAlaGlyArgValThrAspLysTyrPArgAsnPheMetLysGlnIleGlnArgAla 360  
Db 643 GAGCAGCGGCGCGAGAAC-----GGAAGGTGACGCTCGGCTCAACGCGCG 690  
Qy 361 -----ArgLysPhePheAspGluSerGluGlyValThr 372  
Db 691 CTGGCACACTTCATTCGACGTGACGCGGCGGCGGTATTCGTCGCGAATCCGCGATCCG 750  
Qy 373 GluLeuAspSerAlaSerArgTyrProValLeuThrAlaLeuLeuLeuTyrArgLysIle 392  
Db 751 ATGCTCGATTCGCGGCGGTTCGAGCGCTGATCCGACGCGCTTTCGTTGTATCGAGCAATT 810  
Qy 393 LeuAspGluIleGluAlaAsnAspTyrAsnAsnPheThrArgArgAlaTyrVal 410  
Db 811 CTGACACAGGTGAGGCGCGCGCTTCGAGTCTGCAATCGAGTGTCCGTT 864

RESULT 12  
; Sequence 1, Application US/10041472  
; Publication No. US20020092039A1  
; GENERAL INFORMATION:  
; APPLICANT: Shemake, Christine  
; TITLE OF INVENTION: METHODS FOR PRODUCING CAROTENOID COMPOUNDS AND SPECIALITY OILS IN  
; FILE REFERENCE: 16516, 141  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR FILING DATE: 1997-08-08  
; PRIOR APPLICATION NUMBER: US 08/908758  
; PRIOR FILING DATE: 1996-09-08  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1

LENGTH: 1232  
 TYPE: DNA  
 ORGANISM: Echinia uredovora  
 US-10-041-472-1

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score: 1.37e-33	1232	110	55	162	49	11
Percent Similarity: 43.88%						
Best Local Similarity: 29.25%						
Query Match: 15.44%						

US-09-847-081b-2 (1-440) x US-10-041-472-1 (1-1232)

```

Qy 77 SerArgThnGluLysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAla 96
Db 83 AGCCGTCCTTGAAGGGG-----CAATCCGCCGACGAGGCTTCATTCCGAC 127
Qy 97 Gly-----GluMetThrValSerSerGluLysLysValTyrAspValValLeuLysGln 114
Db 128 GGCCTCAATCATCATGATGATCCACGAAAGAGTC----- 166
Qy 115 AlaAlaLeuValLysArgGlnLeuArgSerThrAspAspLeuGluValLysProAsp 133
Db 167 -----AACCTGACATTACTTCATTACCAAGCAATGAT 199
Qy 134 -----IleValValProGluLysLeuGlyLeuLeuSerGluAlaTyrAspArgGly 151
Db 200 GGAGAGATTAAGTGCATGATATATCGTCGTTACTCAATCATGCGGTC----- 247
Qy 152 GluValCysAlaGluTyrAlaLysThrPheTyrLeuGlyThrLysLeuMetThrProGlu 171
Db 248 GAAAGAGTGCAGTGGCTGCAGAAAGTTTGGACAGCCTCAAGTTATTGATGCAAAA 307
Qy 172 ArgArgArgAlaIleTyrAlaIleTyrValTyrCysArgArgThrAspGluLeuValAsp 191
Db 308 ACCCGGCGAGGCTACTGATGCTTACGCGCTGCGCCGCTGATGACGATGTATTGAC 367
Qy 192 GlyProAsn-----AlaSerHisIleThrProGluAlaLeu 203
Db 368 GATCAAGCCGCTGCTTACGCGCGGACGCTGCTTACAAACCCGACAAACGCTG 427
Qy 204 AspArgTyrGluThrArgLeuGluLysIlePheSerGly---ArgProPheAspMetLeu 222
Db 428 ATGCAACTTGATGAATAAAGCGCCAGCGCTTACGATGCCAGATGACAAACCGCG 487
Qy 223 AspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArg 241
Db 488 TTGGCGGCTTTACAGAAAGTGGCTATGCTCAT-----GATATCGCCCGCTTACGCG 541
Qy 242 ---AspMetIleGluGlyMetArgMetAspLeuTyrPheSerArgTyrLysThrPheAsp 260
Db 542 TTGTGATCATCTGAAGGCTTCCGCAATGATGATGACGCAATACGCAACTGAT 601
Qy 261 GluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProVal 280
Db 602 GATAGCGCTGCGCTATTTGCTATCATCGTGCAGCGCTTGTGCTTATGATGCGCAATC 661
Qy 281 MetGlyIleAlaProGluSerLysAlaThrThrGlnSerValTyrAsnAlaAlaLeuAla 300
Db 662 ATGGGCGTG-----CGGATTAACGCCACGCTGAC-----CGGCGCTGTGAC 703
Qy 301 LeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuAlaGspValGlyGluAspAlaArgArg 320
Db 704 CTGGGCTGCGCATTTTCAGTTGACCAATATTGCTCGGATATTGTCGACATCGCATGCG 763
Qy 321 GlyArgValTyrLeuProGluLysAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIle 340
Db 764 GGCCTGTGTTATCTGCGCGCAAGCTGCGTGCAGCATGAAGGCTTGAACAAAGATTTAT 823
Qy 341 PheAlaGlyArgValThrAspLysTyrArgAsnPheMetLysGlnIleGlnArgAla 360

```

```

Db 824 GCGGACCTGAAACCCTCAGCGCGTGAAGCCGTAATGCCCGCTGTTGGTGCAGAAACA 883
Qy 361 ArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTyr 380
Db 884 GAACCTTACTTATTTGTCTGACACAGCGGCTGCGACAGGTTGCCCGCTCCGCTCG 943
Qy 381 ProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAsp 400
Db 944 GCAATGCTACGCGGAGACGAGGTTTACCGGAAATAGTGTCAAAGTTGAACAGCGCGT 1003
Qy 401 TyrAsnAsnPheThrArgAlaGlyAlaTyrValSerLysProLysLysLeuThrLeuPro 420
Db 1004 CAGCAAGCTTGGATGACGCGGACGTCACAGCACCGCCGCAAAATTAAGCTGCTGCTG 1063
Qy 421 IleAlaTyrAlaLysSerLeuValProProAsnArgThrSerSerPro 436
Db 1064 GCCGCTCTGTGTCAGGCTTACTTCCCGATGCGGCTCATCTCTCC 1111

```

## RESULT 13

```

US-09-941-947A-33
; Sequence 33, Application US/09941947A
; Publication No. US20030003528A1
; GENERAL INFORMATION:
; APPLICANT: Brzostowicz, Patricia C.
; APPLICANT: Cheng, Qiong
; APPLICANT: DiCosimo, Deana J.
; APPLICANT: Koffas, Matheos
; APPLICANT: Miller, Edward S. Jr.
; APPLICANT: Odum, J. Martin
; APPLICANT: Picataggio, Steve
; APPLICANT: Rouviere, Pierre E.
; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
; FILE REFERENCE: CL1903 US NA
; CURRENT APPLICATION NUMBER: US/09/941,947A
; PRIOR FILING DATE: 2001-09-01
; PRIOR APPLICATION NUMBER: 60/229,907
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 33
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Pantoea stewartii
US-09-941-947A-33

```

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score: 4.77e-32	891	88	52	131	19	6
Percent Similarity: 48.28%						
Best Local Similarity: 30.34%						
Query Match: 14.82%						

US-09-847-081b-2 (1-440) x US-09-941-947A-33 (1-891)

```

Qy 158 AlaLysThrPheTyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTyr 177
Db 13 TCGAAGAGCTTTGGACGTCATGACGCTTTTCAGCCCAAAACCGGTCGAGGTGCTG 72
Qy 178 AlaIleTyrValTyrCysArgArgThrAspGluLeuValAspGlyProAsn----- 194
Db 73 ATGCTTTACGATGATGTCGCGGACATGCGACGACTCATTTGACATCAAAACGCTGGCTT 132
Qy 195 -----AlaSerHisIleThrProGluAlaLeuAspArgTyrGluThrArg 209
Db 133 CATGCCACGACCCCTCTTGCAGATGCTGACGACGCGCTGACGACGCTTGAATGAA 192
Qy 210 LeuGluAspIlePheSerGly---ArgProPheAspMetLeuAspAlaAlaLeuSerAsp 228
Db 193 ACGGCTCAGGCTTACGCGGCTTGCAGCAATGACGAGCCGCTTGTCCGCGTTTCAGGAG 252

```

```

QY      229 ThrValSerArgPheProValAspIleGlnProPheArg-----AspMetIleGluGly 246
Db      253 GTGCGCATGCGCGCAT-----GATATCGCTCCCGCTACGCGCTTCACCATCATCGAAGGT 306

QY      247 MetArgMetAspLeuTrpLysSerArgTyrLysThrPheAspGluLeuTyrLeuTyrCys 266
Db      307 TTGGCCATGAGTGTGGCGGAACGCCCTACCTGACACTGACGATACGCTGCGCTTATTTCC 366

QY      267 TyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGlyIleAlaProGlu 286
Db      367 TATCAGTGGCGGCTTTGTGGCGCTGATGATGCGCCAAATTATGGCGCTT-----CGC 420

QY      287 SerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGlyLeuAlaAsnGln 306
Db      421 GATTAAGCCACGCTCAT-----CGCGCTGCGATCTCGGCTGGCTTCCAG 468

QY      307 LeuThrAsnIleLeuArgPheValGlyIleAspAlaArgArgGlyValTyrLeuPro 326
Db      469 TTGACCAACATTGGCGGTATTTGTGACGATGCTCAGGTGGCGGCTGTATCTGCCT 528

QY      327 GlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAlaGlyArgValThr 346
Db      529 GAAAGCTGGCTGGAAGAGAGAGACTGACGAAAGCAATTATGCTGCCAGAAACCGG 588

QY      347 AspLysTrpArgAsnPheMetLysGlnIleGlnArgAlaArgLysPhePheAspGlu 366
Db      589 CAGGCTTAAGCCGTATCGCGCGGCGATGCTGAGGAGCGGAACCCATTATCGATCA 648

QY      367 SerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpProValLeuThrAlaLeu 386
Db      649 TCAATGCGCGGCTGGCACAAATTACCTTACGCTCGGCGCTTCGCCAGAGGAAG 708

QY      387 LeuLeuTyrArgLysIleLeuAspGluIleGlnAlaAsnAspTyrAsnAsnPheThrArg 406
Db      709 CAGGTATCCGTAATAATTGGCGGTGAAGTTGAACAGCGCGGTACAGCGCTGGATCAT 768

QY      407 ArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIleAlaTyrAlaLysSer 426
Db      769 CCCAGTCCACGCTCCACCGCGCAAAATTATACGCTTTGCTGACGCGCATCCGGTCAAGCA 828

QY      427 LeuValProProAsnArgThrSerSerPro 436
Db      829 GTTACTTCCCGGATGAAGCATATCCACC 858

RESULT 14
US-10-218-118-9
; Sequence 9, Application US/10218118
; Publication No. US20030148319A1
; GENERAL INFORMATION:
; APPLICANT: Brzostowicz, Patricia
; APPLICANT: Rouviere, Pierre
; APPLICANT: Picataggio, Stephen
; APPLICANT: Cheng, Qiong
; TITLE OF INVENTION: Genes Encoding Carotenoid Compounds
; FILE REFERENCE: CL1876 US NA
; CURRENT APPLICATION NUMBER: US/10/218,118
; PRIOR FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 60/312,646
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Pantoea Stewartii
US-10-218-118-9

Alignment Scores:
Pred. No.: 4, 77e-32 Length: 891
Score: 336.50 Matches: 88
Percent Similarity: 48.28% Conservative: 52
Best Local Similarity: 30.34% Mismatches: 131
Query Match: 14.82% Indels: 19

```

```

DB:      13 Gaps: 6
US-09-847-081b-2 (1-440) x US-10-218-118-9 (1-891)

QY      158 AlaLysThrPheTyrLeuGlyThrLysLeuMetThrProGluAlaArgAlaIleTrp 177
Db      13 TCGAAACCTTTGCGACATCGACATCGACCTTTTTCAGCGCAAAACCGCTGCGCGTGGC 72

QY      178 AlaIleTyrValTrpCysArgArgThrAspGluLeuValAspGlyProAsn----- 194
Db      73 ATGCTTACGATGAGTGGCGGCACCTGCGACGACGATCATGACATCAAACTGGGCTTT 132

QY      195 -----AlaSerHisIleThrProGlnAlaLeuAspArgTrpGluThrArg 209
Db      133 CATGCCAGACAGCCCTTTTCGACATGCTGACAGCGCGCTTCACACGTTGAATGAAG 192

QY      210 LeuGluAspIlePheSerGly---ArgProPheAspMetLeuAspAlaLeuSerAsp 228
Db      193 ACGCTCAGCGCTACGCGGTTCCGCAAAATGACAGACCGCGCTTTTCCGCTTCAGAG 252

QY      229 ThrValSerArgPheProValAspIleGlnProPheArg-----AspMetIleGluGly 246
Db      253 GTGCGCATGCGCGCAT-----GATATCGCTCCCGCTACGCGCTTCACCATCTGAAGGT 306

QY      247 MetArgMetAspLeuTrpLysSerArgTyrLysThrPheAspGluLeuTyrLeuTyrCys 266
Db      307 TTGGCCATGAGTGTGGCGGAACGCCCTACCTGACACTGACGATACCTGCGCTTATTC 366

QY      267 TyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGlyIleAlaProGlu 286
Db      367 TATCAGTGGCGGCTGTGTGGCGCTGATGATGCGCAAAATTATGGCGTT-----CGC 420

QY      287 SerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGlyLeuAlaAsnGln 306
Db      421 GATTAAGCCACGCTCAT-----CGCGCTGCGATCTCGGCTGGCTTCCAG 468

QY      307 LeuThrAsnIleLeuArgPheValGlyIleAspAlaArgArgGlyValTyrLeuPro 326
Db      469 TTGACCAACATTGGCGGTATTTGTGACGATGCTCAGGTGGCGGCTGTATCTGCCT 528

QY      327 GlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAlaGlyArgValThr 346
Db      529 GAAAGCTGGCTGGAAGAGAGAGACTGACGAAAGCAATTATGCTGCCAGAAACCGG 588

QY      347 AspLysTrpArgAsnPheMetLysGlnIleGlnArgAlaArgLysPhePheAspGlu 366
Db      589 CAGGCTTAAGCCGTATCGCGCGGCGACTGATCGGGAAGCGGAACCTATTATCGATCA 648

QY      367 SerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpProValLeuThrAlaLeu 386
Db      649 TCAATGCGCGGCTGGCACAAATTACCTTACGCTTGGCGCTGCGCATCGACAGCGGAAG 708

QY      387 LeuLeuTyrArgLysIleLeuAspGluIleGlnAlaAsnAspTyrAsnAsnPheThrArg 406
Db      709 CAGGTATCCGTAATAATTGGCGGTGAAGTTGAACAGCGCGGTACAGCGCTGGATCAT 768

QY      407 ArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIleAlaTyrAlaLysSer 426
Db      769 CGCCAGTCCACGCTCCACCGCGCAAAATTATACGCTTTGCTGACGCGCATCCGGTCAAGCA 828

QY      427 LeuValProProAsnArgThrSerSerPro 436
Db      829 GTTACTTCCCGGATGAAGCATATCCACC 858

RESULT 15
US-10-166-037-4
; Sequence 4, Application US/10166037
; Publication No. US2003008737A1
; GENERAL INFORMATION:
; APPLICANT: GIRAUD, Eric
; APPLICANT: HANNIBAL, Laure
; TITLE OF INVENTION: Isolated Carotenoid biosynthesis gene cluster involved
in canthaxanthin production and application thereof

```

```

; FILE REFERENCE: 1721-55
; CURRENT APPLICATION NUMBER: US/10/166,037
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: 60/297,272
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1008
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: crtb
US-10-166-037-4

Alignment Scores:
Score: 7,81e-31 Length: 1008
Percent Similarity: 327.50 Matches: 99
Best Local Similarity: 45.68% Conservative: 49
Query Match: 30.56% Mismatches: 137
DB: 15 Gaps: 39

US-09-847-081B-2 (1-440) x US-10-166-037-4 (1-1008)

QY 139 AsnLeuGlyLeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAla 158
DB 22 CATCTGGGGCGGCTGAGCGAGCCCTCATCCCGACGCGC-----TCG 63

QY 159 LysThrPheTyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTyrAla 178
DB 64 AAGAGTTTCGACGCGCATCCAGCTGTCGATCCCGCACCCGTCGACGTCGACCTG 123

QY 179 IleTyrValTyrCysArgArgThrAspGluLeuValAspGly----- 192
DB 124 CTCTACGCTGCTGCTGCGATTGCGACGATCGACGACGACGACGATCTCGAATTGCG 183

QY 193 -----ProAsnAlaSerHisIleThrProGluAlaLeu---AspArgTyr 206
DB 184 CAGGCGCGTGGCGGCGCTGCGCCGACGATCGGACCTTTCGACATGCTCGCGACGACG 243

QY 207 GluThrArgLeuGluAspIlePheSerGlyArgProPhe---AspMetLeuAspAla 225
DB 244 GCGCAGCGCGTGGAG-----GGGCGCGCATGCGCGATCGGATTCGAGGGA 291

QY 226 LeuSerAspThrValSerArgPheProValAspIleGlnProPheArgAspMetIleGlu 245
DB 292 TTGACGCGCGTGTGACGAGACGCGGATTCGACACATCAGCTGTCGAGCTCTGAC 351

QY 246 GlyMetArgMetAspLeuTyrLysSerArgTyrLysThrPheAspGluLeuTyrLeuTyr 265
DB 352 GCGTTCGCGCATGATGTCGACGCGCGCATTCGACGCGCATGACGCGATGACGCTGAC 411

QY 266 CysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGlyIleAlaPro 285
DB 412 TGCTATCAAGTGGCGCGGTGGTGGCGGATGATGTCGCGCATCATGAGC----- 462

QY 286 GluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGlyLeuAlaAsn 305
DB 463 -----GCTCGGAGAGGCGCGCTGAGACCGCGCGCGCATCTCGGCATCTCGCGCTG 513

QY 306 GluLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGlyArgValTyrLeu 325
DB 514 CAGCTCACCAACATGCGCGGTGATCGAGATGCCGACGCGCGCATGATCTG 573

QY 326 ProGluAspGluLeuAlaGlyLeuSerAspGluAspIlePheAlaGlyArgVal 345
DB 574 CCGCAGCATGCTGTGCGAGCGCGCGCGCGCGCGCGAGTTCGCGGAACCGCAGCAT 633

QY 346 ThrAspLysTyrPheArgAsnMetLysGlnIleGlnArgAlaArgLysPheAsp 365
DB 634 CCGCAGCGCGTGTCTGTCGCGCGGCTGTCGATGTCGCGAGCAGTTTACGAG 693

```

```

QY 366 GluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTyrProValLeuThrAla 385
DB 694 GCCAGCGAGCGAGCATCGCCGCGCTCGGTCGCTGCGCGCGGTGAGACGCGCC 753

QY 386 LeuLeuLeuTyrArgLysIleLeuAspGluIleGlnAlaAsnAspTyrAsnAsnPheThr 405
DB 754 CGCGTCTCATCGCCGATCGGCGCGAGGTGATGAGAGCGCGGTCCCGCGCGCTGGGAC 813

QY 406 ArgArgAlaTyrValSerLysProLysLysLeu-----LeuThr 418
DB 814 GCCCGCATCGCCAGACAGCGCGCGCAGAACTCGCGCCGATCGCGCCGACGCGATTGACG 873

QY 419 LeuProIleAlaTyrAlaLysSerLeuValProPro-----AsnArgThrSerSer 435
DB 874 CTGATCTCACCCGACCTGCGGTAAAGTGTCCGCGCGAGCACAACTCTGAGACCGCG 933

QY 436 ProLeuAlaLys 439
DB 934 CCGAAGCGCGCG 945

```

Search completed: January 16, 2004, 13:19:13  
 Job time : 11694 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 16, 2004, 05:38:11 ; Search time 2815 Seconds  
(without alignments)  
3798.923 Million cell updates/sec

Title: US-09-847-081b-2  
Perfect score: 2270  
Sequence: 1 MSMSVALIMVVSPTSEVSNNG.....IYAKSLVPPNRTSSPLAKT 440

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n\_model -DEV=tlh  
-Q=/cgn2.1/USPTO\_spool/US09847081/runat.15012004.154337.1867/app\_query.fasta.1.583  
-DB=EST -QFMT=fastlap -SUFIX=fst -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09847081.@CGN 1.1.2810 @runat.15012004.154337.1867 -NCPU=6 -ICPU=3  
-NMAP -LARGEBUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hic.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hic.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pin.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_png.\*  
27: em\_gss\_vr1.\*  
28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214.5	53.5	819	13	B0636463
2	1211	53.3	1655	11	AY111032
3	1202	53.0	787	12	BM410846
4	1201.5	52.9	1201	11	AY108547
5	1179	51.9	781	13	B0511016
6	1136	50.0	751	12	B0046203
7	1130	49.8	765	12	BM408984
8	1099	48.4	686	9	AM442101
9	1083	47.7	682	12	BM409200
10	1072	47.2	672	9	AM442407
11	1049	46.2	642	10	BE433198
12	1043	45.9	676	9	AM222027
13	1001	44.1	677	10	BG351357
14	997	43.9	634	12	BM412533
15	984	43.3	704	14	CA999095
16	980	43.2	600	10	BE460889
17	977.5	43.1	623	10	BE432511
18	956	42.1	685	9	AM221932
19	947	41.7	581	9	AM223316
20	940	41.4	581	10	BE432955
21	939	41.4	623	12	BM536249
22	934	41.1	582	9	AM223666
23	927	40.8	605	10	BE435064
24	927	40.8	744	12	BM412719
25	921	40.6	562	10	BE433966
26	918	40.4	634	12	BM137086
27	917	40.4	565	10	BE432595
28	916	40.4	573	10	BE435308
29	915	40.3	881	14	CB621108
30	913	40.2	585	9	AM222245
31	907	40.0	668	12	B1955682
32	905	39.9	869	14	CB622557
33	899	39.6	556	9	AM222152
34	897	39.5	609	10	BE322877
35	891	39.3	581	9	AM222806
36	889	39.2	690	14	CB342702
37	884	38.9	791	9	AM496851
38	884	38.9	645	9	AM223528
39	881	38.8	542	10	BE431550
40	879	38.7	613	12	B1957464
41	877	38.6	543	12	BM411875
42	873	38.5	634	9	AM441216
43	870	38.3	542	10	BE434578
44	869	38.3	542	10	BE432748
45	869	38.3	888	14	CB621110

## ALIGNMENTS

RESULT 1  
B0636463  
LOCUS B0636463  
DEFINITION B0636463  
ACCESSION B0636463  
VERSION B0636463  
KEYWORDS B0636463.1 GI:23303718  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroside II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 819)  
 Lundegard, M., Emmersen, J., Nielsen, K.L., Wilson, I., Somerville, S.  
 AUTHORS  
 TITLE EST sequencing of Erysiphe cichoracearum infected Arabidopsis  
 plants  
 JOURNAL Unpublished  
 COMMENT Contact: Karen G. Welinder  
 Institut for bioteknologi  
 Aalborg Universitet  
 Sohngaardsholmsvej 49, 9000 Aalborg, Denmark  
 Tel: +45 96358467  
 Fax: +45 98141808  
 Email: kgw@bio.auc.dk.

FEATURES  
 source 1..819  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Columbia"  
 /db\_xref="taxon:3702"  
 /dev\_stage="Plant 3 weeks old, three days post infection"  
 /clone\_lib="Infected Arabidopsis leaf"  
 /note="Organ: Leaf; Vector: pBluescript; Mixed cDNA  
 library of Arabidopsis and E. cichoracearum infected leaf  
 from three weeks old Arabidopsis plants. Plants were  
 harvested 3 days after infection and mRNA oligo dt  
 selected."

BASE COUNT 245 a 158 c 205 g 211 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 5.78e-130 Length: 819  
 Score: 1214.50 Matches: 220  
 Percent Similarity: 92.99% Conservative: 22  
 Best Local Similarity: 84.87% Mismatches: 16  
 Query Match: 53.50% Indels: 3  
 Gaps: 1

US-09-847-081b-2 (1-440) x BU636463 (1-819)

Oy 165 ThrTysLeuMetThrProGluValArgArgAlaIleTTPAlaIleTTPValTTPCyArg 184  
 |||  
 Db 2 ACTTGGCTTATGACACCGAAAGGGGAAAGCGATTGGCAATCTACGTTGGTGTAGA 61  
 |||

Oy 185 ArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGlnAlaLeuAsp 204  
 |||  
 Db 62 AGAAGCTGATGAACCTTGTGATGGCCAAATGCTTACATATATCTCCATGCTTATAT 121  
 |||

Oy 205 ArgTTPGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeuAspAla 224  
 |||  
 Db 122 AGATGGGAAGCAAGCTTAGAAGATCTTTCCGTCGTCCTTCGATATGCTTATGATGCT 181  
 |||

Oy 225 AlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArgAspMetIle 244  
 |||  
 Db 182 GCTTCGCTGATACAGTGTAGATATACCCGGTCGATATTCAGCCATTCGAGACATGATC 241  
 |||

Oy 245 GluGlyMetArgMetAspLeuTrpLysSerArgTyrLysThrPheAspGluLeuTyrIleu 264  
 |||  
 Db 242 GAAGGAATGAGATGACTTGAAGAAATCGAGATACCGAACTGATGATCTATACCTT 301  
 |||

Oy 265 TyrCyTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGlyIleAla 284  
 |||  
 Db 302 TACTGCTACTACGTGCTGGAACCGTCGATGATGACGTTCCGTTATATGGAAATCAT 361  
 |||

Oy 285 ProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGlyLeuAla 304  
 |||  
 Db 362 CCTATGTGAAGCAACCAACCGAAGTGTATTAACAACGTCCTTGCCCTTGATATACC 421  
 |||

Oy 305 AsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGlyArgValTyr 324  
 |||  
 Db 422 AATCAGCTTACTATACATCTCGAGACGTTAGCGAAGATGCGAAGAGAGAGAGGCTTAT 481  
 |||

Oy 325 LeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAlaGlyArg 344  
 |||

Db 482 CTGCTCAGATGAAATTTGGCTCAGCTGTCTTTTCAGATGAGAAGACATATTCGCCGAAAA 541  
 |||  
 Oy 345 ValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnAlaArgAlaArgLysPhePhe 364  
 |||  
 Db 542 GTAACTGATTAATGAGAACTTATCATGAAATGACGCTTAACAGCAAGAAATTTCTTC 601  
 |||

Oy 365 AspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTTPProValLeuThr 384  
 |||  
 Db 602 GACCAAGCTGGAAGAGCGTCACCGAGCTCAGTCGCCCTAGCAGATGGCCTGTATGGGCT 661  
 |||

Oy 385 AlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGlnAlaAsnAspTyrAsnAspPhe 404  
 |||  
 Db 662 TCATTTGCTATTGTACAGAGAAATCTGACGAGATTGAAGCAATGATTAACAACAATTT 721  
 |||

Oy 405 ThrArgArgAlaTyrValSerLysPheLysLeuLeuThrLeuProIleAlaTyrAla 424  
 |||  
 Db 722 ACTAAGAGCTTATGTGGGAAAGTCAAGAAATTCAGCTTTCGATTCGCTTATGCT 781  
 |||

Oy 425 LysSerLeuValProProAsnArgThrSerSer 435  
 |||  
 Db 782 AATCAGTACTA-----AAGACTTCAAGT 805  
 |||

RESULT 2  
 AV111032  
 LOCUS  
 DEFINITION Zea mays CL1906\_1 mRNA sequence.  
 ACCESSION AV111032  
 VERSION AV111032.1 GI:21215622  
 KEYWORDS  
 HTCC  
 ORGANISM  
 Zea mays  
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoidae; Andropogoneae; Zea.  
 1 (bases 1 to 1655)  
 Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S.,  
 Arthur, L.W., Hanafey, M., Morgante, M., and Tingey, S.V.  
 Maize Mapping Project/Dupont Consensus Sequences for Design of  
 Overgo Probes  
 Unpublished (2002)  
 2 (bases 1 to 1655)  
 Coe, E.H.  
 Direct Submission  
 Submitted (25-APR-2002) Maize Mapping Project, University of  
 Missouri, Columbia, MO 65211, USA  
 If you are interested in getting corresponding physical clones,  
 these are publicly available from ZmDB and may be found by BLAST  
 searching at MSU, [maizemap.org](http://maizemap.org); ZmDB, [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu); TIGR,  
[www.tigr.org](http://www.tigr.org); or NCBI, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov). When the source of the  
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
 Schnable, Iowa State, then clones may be requested from ZmDB:  
[www.zmdb.iastate.edu](http://www.zmdb.iastate.edu).

FEATURES  
 source 1..1655  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4577"  
 /clone\_lib="Maize Mapping Project/Dupont Consensus  
 Library"  
 /note="this sequence is part of a project of EST  
 assemblies resulting from the application of public  
 contigs to seed Dupont contigs; this resource was  
 assembled by Dupont as part of a collaboration for the  
 overgo addressing of BACs in conjunction with the Maize  
 Mapping Project"

BASE COUNT 384 a 334 c 380 g 295 t 262 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 4.23e-129 Length: 1655  
 Score: 1211.00 Matches: 244  
 Percent Similarity: 79.47% Conservative: 27





FEATURES	LOCATION/Qualifiers
Db	127 ACCAATGAGTTAGAAAGTGAAGCCGGAATTAACCTATTCCGGGGAAATTGGGCTGTGTAGT 186
Oy	145 GUAUATATyAspAspArgCyGgIyGluValCyAlaGluTyAlaAlaTyTh-PhetYrLeuGly 164
Db	187 GAAGCATATGATGATAGGCTGTGTGTAAGATATGTCGACAGATATGCMAAACGTTTAACCTTAGA 246
Oy	165 ThrLysLeuMetTh-ProGluArgArgArgAlaIleTETPAlaIleTyValTProCyArg 184
Db	247 ACTATGCTTAATGACTCCGAGAGAAAGAGGCTATCTGGGCAATATATGATGTGGCGAGA 306
Oy	185 ArgThrAspGluLeuValAspGlyProAlaAlaSerHisIleTh-ProGluAlaLeuAsp 204
Db	307 AGAACAATGATGAACCTTGTTGATGGCCCAAAACGATCATATATTATACCOCGGACCTTAAT 366
Oy	205 ArgTTPGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeuAspAla 224
Db	367 AGGTGGGAAAATAGGCTAGAGATGTTCATATGGGCGGCCATTTGACATGCTCCATGAT 426
Oy	225 AlaLeuSerAspTh-ValSerArgPheProValAspIleGlnProPheArgAspMetIle 244
Db	427 GCTTTGTCGACATACGTTTCTTAACCTTCCAGTTGATATTCAGCCATTCAAGATATGAT 486
Oy	245 GluGlyMetCArgMetCAspLeuTrpLysSerArgTyIYrThrPheAspGluLeuTyrLeu 264
Db	487 GAAGGAATGCCGATATGACTCTGAGAAATATGACATATCAAAAATCTTGACAGCACTATACCT 546
Oy	285 TyrCySTyTyTyValAlaGlyThrValGlyLeuMetSerValProValMetGlyIleAla 284
Db	547 TATTTTATTTATTTATGTTGCTGTGATGCGTTGGGTTGATGAGAGTTTCCAATTATGGGTATCC 606
Oy	285 ProGluSerIyAlaTh-ThrGluSerValTyArgAlaAlaLeuAlaLeuGlyLeuAla 304
Db	607 CCTGATCAAAAGCAACAACAGAGAGCGATATATATGCTGCTTGCTGGGATGCCA 666
Oy	305 AsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGlyArgValTy 324
Db	667 AATCAATTATTAACATCTACTCAGAGATGTGGAGAAAGATCCAGAAAGAAAGATCTAC 726
Oy	325 LeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAlaGlyArg 344
Db	727 TTGGCCTCAGATGAATTATGACACAGGCAAGGCTATCCGATGAAGATATATTCTCGAAG 786
RESULT 4	
LOCUS	AY108547 1201 bp mRNA linear HTC 17-OCT-2002
DEFINITION	Zea mays PCO131047 mRNA sequence.
ACCESSION	AY108547
VERSION	AY108547.1 GI:21211629
KEYWORDS	HTC.
SOURCE	Zea mays
ORGANISM	Zea mays
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
AUTHORS	1 (bases 1 to 1201) Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
JOURNAL	Unpublished (2002)
REFERENCE	2 (bases 1 to 1201) Coe,E.H.
AUTHORS	Direct Submission
TITLE	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
JOURNAL	Misouri, Columbia, MO 65211, USA
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDb and may be found by BLAST searching at MSU, maizegap.org; ZmDb, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDb: www.zmdb.iastate.edu.

Source	1. 1201	/organism="Zea mays"	/mol_type="mRNA"	/db_xref="taxon:4577"	/clone_lib="Maize Mapping Project/Dupont Cornsensus library"	/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"
BASE COUNT	273 a	325 c	316 g	287 t		
ORIGIN						
Alignment Scores:						
Pred. No.:	3.3e-128	Length:	1201			
Score:	1201.50	Matches:	223			
Percent Similarity:	92.62%	Conservative:	28			
Best Local Similarity:	82.29%	Mismatch:	19			
Query Match:	52.93%	Indels:	1			
DB:	11	Gaps:	1			
US-09-847-081B-2 (1-440) x AY108547 (1-1201)						
QY	161 PheTyrIleu--GlyThrIysLeuMetThrProGluIarGatGatGAlaIleTTPaIle	179				
DB	167 TTTTACGCTTTCAGGCGACGACGCTCATGACTCTCGAGCGGCCCAAGCGCTCGGCGATC	226				
QY	180 TyrValTTCyAaGrgArGhrzAspGluLeuValAspGlyProAnaIaSerHisIleThr	139				
DB	227 TAGGTGGGAGCAAGAACTACAGACTAGAGAGGTGCCAACAGGCTCTACATACG	286				
QY	200 ProGlnAlaLeuAspArgTgTgPgluThrArgLeuGluAspIlePheSerGlyArgProPhe	219				
DB	287 CCGACCGCTCTGCACCGCGCTGGAGAGAGCGGCTGAGAGATCTCTTGAGGGCGCGCGTAC	346				
QY	220 AspMetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnPro	239				
DB	347 GACATGACGACGCGCGCTCTCGACACTGTGTCCAGATTCCCGCTCATACACCG	406				
QY	240 PheArgAspMetIleGluGlyMetArgMetAspLeuTplySerArgTyrIysThrPhe	255				
DB	407 TTCAAGACATGTGTCACAGAAATGAGCGCTGACCTGTGAAGTGCAGATACATGACCTTC	466				
QY	260 AspGluLeuTyrIleuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValPro	279				
DB	467 GACGAGCTCACTCTCACTGCTCACTACGTCCGCGGACCGCTCGGCTCATGACGTCCT	526				
QY	280 ValMetGlyIleAlaProGluSerIysAlaThrThrGluSerValTyrAsnAlaLeu	299				
DB	527 GTCATGGGATGCTCCGACCTCCAAGGCGTGCACCGACGCTCAATGCTGCTCTG	586				
QY	300 AlaLeuGlyLeuAlaAsnGluLeuThrAsnIleLeuArgAspValGlyGluAspAlaArg	319				
DB	587 GCTCTCGGCACTCCCTAACACGCTGCAGATATCTCCACAGACGCGGCGAAGAGCGAGG	646				
QY	320 ArgGlyArgValTyrIleuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAsp	339				
DB	647 AGGGGAGAAATATACCTTCCTTGACACGACTGCGACGACGCTCTACCGAAGAGAGAC	706				
QY	340 IlePheAlaGlyArgValThrAspIysTTPArgAspMetIysGlnIleGlnArg	355				
DB	707 ATATTTCAGAGGAAAGTGAACGCAATGAGAGATTCATGAAGGCGACATCCACGCT	766				
QY	360 AlaArgIysPhePheAspGluSerGlyIysGlyValThrGluLeuAspSerAlaSerArg	379				
DB	767 GCCAGGCTCTCTTTATAGAGCGGAGAGAGGCGTCAACCTTCGACTGCTACGAGA	826				
QY	380 ThrProValLeuThrAlaLeuLeuTyrArgIysIleLeuAspGluIleGluAlaAsn	399				
DB	827 TGGCCGATGCTCCGCTCTGTGGCTGATACGACGATCTTGATGCCATTGAGGCAAC	886				

QY 400 APTTYAsnaSnpheThirgaArgAlaTYValSerLySProLySLeuLeuThrLeu 419  
 |||.....  
 DB 887 GACTACAAACAATTCCACCAAGGCTGGTACGTGGCAAGGCCAAGAAGCTGCTGTGTTA 946  
 QY 420 ProLleAlATYrAlaLySLeuValProPro 430  
 |||.....  
 DB 947 CCGCTTCATATGCACAGGCTGCTTGACCA 979

RESULT 5  
 B0511016/c  
 LOCUS B0511016 781 bp mRNA linear EST 07-MAR-2003  
 DEFINITION EST18431 Generation of a set of potato CDNA clones for microarray  
 analyses mixed potato tissues Solanum tuberosum CDNA clone STMH067  
 3' end, mRNA sequence.

ACCESSION B0511016  
 VERSION B0511016.2 GI:21926690  
 KEYWORDS EST.  
 SOURCE Solanum tuberosum (potato)  
 ORGANISM Solanum tuberosum  
 Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum.  
 1 (bases 1 to 781)  
 Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,  
 Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and  
 Karamycheva, S.A.  
 Generation of a set of potato CDNA clones for microarray analyses  
 Unpublished  
 On Jun 10, 2002 this sequence version replaced gi:21369885.  
 Other ESTs: EST618430  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@igrr.org  
 This clone can be obtained from the University of Arizona Genomics  
 Institute. Orders can be made through URL:  
 http://genome.arizona.edu/orders/  
 Seq primer: T7.

FEATURES  
 source Location/Qualifiers  
 1..781  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec or Bintje"  
 /db\_xref="taxon:4113"  
 /clone="STMH067"  
 /tissue\_type="mixed tissues"  
 /lab\_host="SOLR"  
 /clone\_lib="Generation of a set of potato CDNA clones for  
 microarray analyses mixed potato tissues"  
 /note="Vector: pBluescript SK(-), Site 1: EcoRI, Site 2:  
 XhoI; supplier: Combination of untreated and Phytophthora  
 infestans-treated libraries of stolons, leaves, leaflets,  
 axillary buds of stem explants, petioles, germinating eyes  
 tubers, or roots."

BASE COUNT 202 a 191 c 144 g 244 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 7.04e-126 Length: 781  
 Score: 1179.00 Matches: 230  
 Percent Similarity: 96.36% Conservative: 8  
 Best Local Similarity: 93.12% Mismatches: 9  
 Query Match: 51.94% Indels: 1  
 DB: 13 Gaps: 0

US-09-847-081B-2 (1-440) x B0511016 (1-781)

QY 194 AsnaLaserHisLlePhProGlnAlaLeuAspArgTrpGluThrArgLeuGluAspIle 213  
 |||.....  
 DB 781 AATGCAATCCACATTAATCCACCAAGCTTATAGAGTGGAGGCCCGCTGGAGAGATAT- 723  
 QY 214 PhSeSerGlyArgProPhSeAspMetLeuAspAlaAlaLeuSerAspThrValSerArgPhe 233

DB 722 TTCACGGGGCGGCCATTGATATGCTTATGTCAGCTTATCCGATATCGTTTCCAATT 663  
 QY 234 ProValAspIleLeuProPhSeArgSpmetIleGluGlyMetArgMetAspLeuTrpLys 253  
 DB 662 CCGTGTGATATTCACCATTCAGATATGCTGTAAGGAATCCGATGACCTGTGGAAA 603  
 QY 254 SerArgTYrLYLeuThPhSeAspGluLeuTYrLeuTYrCYrTYrTYrValAlaGlyThrVal 273  
 DB 602 TCCAGATACACAACTTGTATGATTAATATCTATATGTTACTTATGTCGCTGGACGTA 543  
 QY 274 GlyLeuMetSerValProValMetGlyIleAlaProGluSerLYLeuAlaThrThrGluSer 293  
 DB 542 GGATTGATGAGAGTTCATTATATGAGCACTTGCACTGCAATCAAGGCAACAGACAGAT 483  
 QY 294 ValTYrAsnaAlaAlaLeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAsp 313  
 DB 482 GTATATTAACGCGCTTGGCTTTAGGAGATCGCAAAATCACTAACCAATATCTCAGAGAT 423  
 QY 314 ValGlyGlyAspAlaArgArgGlyArgValTYrLeuProGlnAspGluLeuAlaGlnAla 333  
 DB 422 GTAGAGAAAGATGCAAGAAAGAGATATCTTACCTCAAGATGAAATTAGCACAGCA 363  
 QY 334 GlyLeuSerAspGluAspIlePheAlaGlyArgValThrAspLYLeuTrpArgSpmMet 353  
 DB 362 GGGCTCTCCGATGAAGACATTTTCTGGAAGAGTACGATTAAGTGAGAGATCTTTATG 303  
 QY 354 LysLYSglnIleGlnArgAlaArgLYSPhSeAspGluSerGlyValThrGlu 373  
 DB 302 AAGAAGCAAAATTCACAGGGCAGGAATCTTTGATGAGGCGAAGAAAGTGTCCACAGA 243  
 QY 374 LeuAspSerAlaSerArgTrpProValLeuThrAlaLeuLeuTYrArgLYSLeu 393  
 DB 242 CTGACCTCTGCTAGTAGAAGGCGGAGTGTGGCTGTCTATATCGCAAGATACG 183  
 QY 394 AspGlnIleGluAlaAsnAspTYrAsnaSnpheThrArgArgAlaTYrValSerLYSPro 413  
 DB 182 GACGAGATTGAAGCAAGACGATCAACAACATTCACAGAGGGCTTATGTGCAAGCA 123  
 QY 414 LysLYSLeuLeuThLeuProIleAlaTYrAlaLYSLeuValProPhAsnArgThr 433  
 DB 122 AAGAAGCTTCGACCTTCCATTCCTTATGCAAGATCTCTAGTCCCTTAATGCAACT 63  
 QY 434 SerSerProLeuAlaLYSThr 440  
 DB 62 TCTTCCCACTAGCAAGACA 42

RESULT 6  
 B0046203  
 LOCUS B0046203 751 bp mRNA linear EST 10-MAR-2003  
 DEFINITION EST95321 P. infestans-challenged potato leaf, incompatible  
 reaction Solanum tuberosum CDNA clone BPL14E21 5' end, mRNA  
 sequence.

ACCESSION B0046203  
 VERSION B0046203.1 GI:19820189  
 KEYWORDS EST.  
 SOURCE Solanum tuberosum (potato)  
 ORGANISM Solanum tuberosum  
 Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum.  
 1 (bases 1 to 751)  
 Zhang, P., Hernandez, M., Tornqvist, C.-E., Wirtz, U., Loukoianov, A.,  
 Rangel, O., Haberlach, G.T., Karamycheva, S.A., Tsai, J., Chienlingo, A.,  
 Bougri, O., Buell, C.R., Romling, C.M., Helgeson, J. and Baker, B.  
 Generation of ESTs from Potato Leaves Challenged with Phytophthora  
 infestans, incompatible interaction (2002)  
 Unpublished

TITLE  
 JOURNAL  
 COMMENT  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@igrr.org  
 This clone can be obtained from the University of Arizona Genomics

Institute. Orders can be made through URL:  
http://genome.arizona.edu/orders/  
Seq primer: T3.

FEATURES  
SOURCE Location/Qualifiers

```
1..751
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="BPL114E21"
/tissue_type="leaf"
/dev_stage="6 week old"
/lab_host="SOLR"
/clone_1ib="P. infestans-challenged potato leaf,
incompatible reaction"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: UC Berkeley, PGSC; sequencing: The
Institute for Genomic Research. Whole plants were
challenged with 450,000 sporangia/ml P. infestans isolate
US-1 (US940501) in Biotron (Madison, Wisconsin). Leaf
tissue was collected at 1, 2, 5, 12, and 24 hours
post-challenge and frozen in liquid nitrogen immediately
upon removal. Kennebec plants showed no signs of HR.
Katahdin plants (susceptible to P. infestans US-1) were
used as controls and showed infection. NOTE: We cannot
exclude the possibility that this sequence is actually
derived from Phytophthora rather than potato."
```

BASE COUNT 219 a 119 c 199 g 214 t

ORIGIN

Alignment Scores:

```
Pred. No.: 6.44e-121 Length: 751
Score: 1136.00 Matches: 225
Percent Similarity: 94.40% Conservative: 11
Best Local Similarity: 90.00% Mismatches: 12
Query Match: 50.04% Indels: 3
DB: 12 Gaps: 0
```

US-09-847-081b-2 (1-440) x BQ046203 (1-751)

```
QY 67 AAlaAPPpRoAgtTyrSerCylLeuGlySerAgtThrGluValGlySerThrPheSer 86
DB 5 GCAGATTGAGATATTCGTGTTAGAGATCAACAACTGAGAAATGAGAGAGGCTTTTCT 64
QY 87 ValGlnSerSerLeuValAlaSerProAlaGlyGluMetThrValSerSerGluValAsp 106
DB 65 GTACAGTCCAGTTTGCTGGCTAGTCCAGCTGAGAAATGGCTGTCTATCAGAGAAAG 124
QY 107 ValTyrAspValAlaLeuLysGlnAlaAlaLeuValLysArgGlnLeuArgSerThrAsp 126
DB 125 GTGATATGAGGTGTATTTAGAGAGCAGCTTTAGTGAAGAGCAGCTGATATCTACTGAG 184
QY 127 AspleuGluValLysProAspLleValValProGlyAsnLeuGlyLeuLeuSerGluAla 146
DB 185 GACATAGAGAGAACCCGATATTCGTTCCGGCTAATTTGGGCTTGTGAGTGAACGA 244
QY 147 TyrAspArgCysGluValCysAlaGluTyrAlaLysThrPheTyrLeuGlyTyrLys 166
DB 245 TATGATTCGTGTGGAGAGATATGCAAGATATGCTTAAGACATTTACTTGAAGACCAATG 304
QY 167 LeuMetThrProGluArgArgAlaLleTyrAlaLleTyrValTyrCysArgArgThr 186
DB 305 CTAAATGACTCCAGACAGAAAGAGCTATCTGGCAATATATGTGTGGCGAGGAACT 364
QY 187 AspleuLeuValAspGlyProAsnAlaSerHisLleThrProGlnAlaLeuAspArgTyr 206
DB 365 GATGAGCTTGTGTGATGGCCCTAATCATCACATCACTCCCAAGCTTTGATAGTGGTGG 424
QY 207 GluThrArgLeuGluAspLlePheSerGlyArgProPheAspMetLeuAspAlaLeu 226
DB 425 GAGGCGAGGCTGGAGATATTTTCAACGGCGGCAATTTGATATCTTGAAGACACTTGA 484
QY 227 SerAspThrValSerArgPheProValAspLleGlnProPheArgAspMetLleGluGly 246
```

```
DB 485 TCCGATACGTGTTCCAAATTCCTGTTGATATTCAGCCATTCAGAGATATGTTGAAGA 544
QY 247 MetArgMetAspLeuTyrLysSerArgTyrLysThrPheAspGluLeuTyrLeuTyrCys 266
DB 545 ATGCGTATGAGCTGTGTGGAATCCAGATACACAACTTTGATGAACTATATATTTGT 604
QY 267 TyrTyrValAlaGlyThrValGly-LeuMetSerValProValMetGlyLleAlaProGly 286
DB 605 TACTATGTCGTGTATAGATAGGAATTTGATGAGTTCCAATTAATGAGCATTCGACCTGA 664
QY 286 USeLysAlaThrThrGluSerValTyrAsnAlaAla-LeuAlaLeuGlyLeuAlaAsnG 306
DB 665 ATCCAGAGCAACGACAGAGATGATATTAACGACGCTTTTGGCTTTGAGGATCCCAATC 724
QY 306 LInLeuThrAsnLleLeuArgAspVal 314
DB 725 AACTAACCAT-ATACTCAAGATGTA 749
```

RESULT 7  
LOCUS BM408984 765 bp mRNA linear EST 22-JAN-2002  
DEFINITION EST583311 tomato breaker fruit Lycopersicon esculentum cDNA clone  
C1EG46P23 5' end, mRNA sequence.  
ACCESSION BM408984  
VERSION BM408984.1 GI:18260614  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum

REFERENCE  
AUTHORS Alcala, J., Vrebalov, J., White, R., Vission, T., Karamycheva, S. A., Tsai, J., Bougri, O., Kikines, R., Uterback, T., Van Aken, S., Roming, C. M., Fraser, C. M., Martin, G. B., Tanksley, S. D. and Giovannoni, J.  
TITLE Generation of ESTs from tomato fruit tissue, breaker stage (2002)  
JOURNAL Unpublished  
COMMENT Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: http://www.genome.clemson.edu/orders/index.html  
This clone is available through the Clemson University Genomics Institute.  
Seq primer: T3.

FEATURES  
SOURCE Location/Qualifiers

```
1..765
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="C1EG46P23"
/tissue_type="Pericarp"
/dev_stage="Breaker"
/lab_host="SOLR"
/clone_1ib="tomato breaker fruit"
/note="Vector: pBluescriptSKmudapt; Site_1: EcoRI;
Site_2: XhoI; supplier: Boyce Thompson Institute; Fruit
sequencing: The Institute for Genomic Research. Fruit
were harvested at the breaker stage (first sign of
lycopene accumulation on the blossom end of fruit). Fruit
were cut in half and the seeds and locules were discarded
prior to freezing the pericarp."
```

BASE COUNT 237 a 122 c 202 g 204 t

ORIGIN

Alignment Scores:

```
Pred. No.: 3.29e-120 Length: 765
Score: 1130.00 Matches: 222
Percent Similarity: 90.94% Conservative: 9
Best Local Similarity: 87.40% Mismatches: 20
Query Match: 49.78% Indels: 3
```

DB: 12 Gaps: 2

US-09-847-081b-2 (1-440) x BM408984 (1-765)

QY 128 LeuGluValIysProAspIleValProGlyAsnLeuGlyLeuLeuSerGluAlaTyr 147  
 1 TTAGAAGTGAAGCCGGATATCTATTCGGGGAAATTTGGCTGTGTGAGTGAAGCATAT 60

QY 148 AspArgCysGlyGluValCysAlaGluTyrAlaIleThrPheTyrLeuGlyThrIleu 167  
 61 GATAGGTGGTGAAGTATATGTGACAGTATGCAAAACGCTTTAACTTAGGAAGTATGCTA 120

QY 168 MetThrProGluArgArgAlaIleTrrPalalIeTrrValTrrPyrCysArgArgThrAsp 187  
 121 ATGACTCCCGAGAGAAGAGAGGGCTATCTGGGCAATATATGTATGTGCGAAGAAACAGAT 180

QY 188 GluLeuValAspGlyProAsnAlaSerHisIleThrProGluAlaLeuAspArgTrpGlu 207  
 181 GAACCTGTGTGATGGCCCAACGCAATATATATCCCGGACGCTTAGATAGTGGGAA 240

QY 208 ThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeuAspAlaIleLeuSer 227  
 241 AATAGGCTTGAAGATCTTTCAATGGCGCCGCTTTGACATGCTCGATGGCTTTGTCC 300

QY 228 AspThrValSerArgPheProValAspIleGlnProPheArgAspMetIleGluGlyMet 247  
 301 GATACAGTTTCACTTCCAGTTGATATTCAGCCCATTCAGAGATATGATTTGAAGGAATG 360

QY 248 ArgMetAspLeuTrpIysSerArgTyrIleThrPheAspGluLeuTyrIleuTyrCysTyr 267  
 361 CGTATGGACTTGAGAAATCGAGATCAAAACCTTCGACGAACCTATACCTTATTTGTAT 420

QY 268 TyrValAlaGlyThrValGlyLeuMetSerValProValMetGlyIleAlaProGluSer 287  
 421 TATGTTCGCTGAGACGTTGGTGGTTGATGATGATTCCTCAATTAAGGATTCGCTGATCA 480

QY 288 LysAlaThrThrGluSerValTyrAsnAlaIleLeuAlaLeuGlyLeuAlaAsnGlnLeu 307  
 481 AAGGCACACACAGAGAGCGTATATATATGCTTGGCTTGGGAGTCCCAATCAATTA 540

QY 308 ThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGlyArgValTyrLeuProGln 327  
 541 ACTAACATTACTCAGAAATGTGGAGAAATGTCGAGAAAGAGAAAGATCTTACTTGCCTCA 600

QY 328 AspGluLeuValGluAlaGlyLeuSerAspGluAspIle-PheAlaGlyArgValThrAsp 347  
 601 GATGAATTTGCAAGCAGGAGCTCTATCCAGTGAAGATATTTTGGCTGGAAGGGTGAACCA 660

QY 347 PylsTrpArgAsnPheMetLysLysGlnIleGlnArgAlaArgLysPheAspGluIle 367  
 661 TAAATGGAGAATCTTATATGAAGAAACA---TACTAGGGCAGAAAGTCTTTGATGATGACA 717

QY 367 rGluLysGlyValThrGluLeuAspSerAlaSerArgTrp 380  
 718 GAGAAAGCGCT---GACAGATGAGCTCAGCTAGTATGATTC 754

Db

RESULT 8  
 AM442101

LOCUS EST1197 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA

DEFINITION

ACCESSION AM442101

VERSION AM442101.1 GI:6977352

KEYWORDS EST.

ORGANISM Lycopersicon esculentum (tomato)

SOURCE Lycopersicon esculentum

EUKARYOTA: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; asterids: lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 686)

Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Romming,C.M., Frazer,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.

TITLE Generation of ESTs from tomato fruit tissue

JOURNAL Unpublished

COMMENT Contact: CUGI

Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES

source

1..686

location/Qualifiers

1..686

/organism="Lycopersicon esculentum"

/mol\_type="mRNA"

/culivar="TA496"

/db\_xref="taxon:4081"

/clone="clem21G20"

/tissue\_type="pericarp"

/dev\_stage="red ripe (7-20 days post-breaker)"

/clone\_11b="tomato fruit red ripe, TAMU"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Giovannoni; Fruit were tagged at the blossom end of the fruit) and harvested 7 days post-breaker (fully red-ripe), 10 days post breaker, and 20 days post-breaker (over-ripe). 20 day fruit which showed external or internal signs of pathogenesis were discarded. Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

BASE COUNT 204 a 114 c 182 g 186 t

ORIGIN

Alignment Scores:

Pred. No.: 1..16

Score: 1099.00 Length: 686

Percent Similarity: 95.61% Matches: 209

Best Local Similarity: 91.67% Conservative: 9

Query Match: 48.41% Mismatches: 10

DB: 9 Indels: 0

Gaps: 0

US-09-847-081b-2 (1-440) x AM442101 (1-686)

QY 108 TyrAspValIleLeuLysGlnAlaIleValIysArgGlnLeuArgSerThrAsp 127  
 2 TATGATGTGGTTTGAAGCAGGACGCTTGGTGAAGGCACTGAGATCTACCAATGAG 61

Db

QY 128 LeuGluValIysProAspIleValIleValProGlyAsnLeuGlyLeuLeuSerGluAlaTyr 147  
 62 TTAGAAGTGAAGCCGGATATATCTTCCGGGAATTTGGCTTGTGATGAACATAT 121

QY 148 AspArgCysGlyGluValCysAlaGluTyrAlaIleThrPheTyrLeuGlyThrIleu 167  
 122 GATAGGTGGTGAAGTATATGTGACAGTATGCAAAACGCTTTAACTTAGGAAGTATGCTA 181

QY 168 MetThrProGluArgArgAlaIleTrrPalalIeTrrValTrrPyrCysArgArgThrAsp 187  
 182 ATGACTCCCGAGAGAAGAGGGCTATCTGGGCAATATATGTATGTGCGAAGAAACAGAT 241

QY 188 GluLeuValAspGlyProAsnAlaSerHisIleThrProGluAlaLeuAspArgTrpGlu 207  
 242 GAACCTGTGTGATGGCCCAACGCAATATATATCCCGGACGCTTAGATAGTGGGAA 301

QY 208 ThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeuAspAlaIleLeuSer 227  
 302 AATAGGCTTGAAGATGTTTCAATGGGCGCCATTTGACATGCTCGATGGCTTTGTCC 361

QY 228 AspThrValSerArgPheProValAspIleGlnProPheArgAspMetIleGluGlyMet 247  
 362 GATACAGTTTCACTTCCAGTTGATATTCAGCCCATTCAGAGATATGATTTGAAGGAATG 421

QY 248 ArgMetAspLeuTrpIysSerArgTyrIleThrPheAspGluLeuTyrIleuTyrCysTyr 267  
 422 CGTATGGACTTGAGAAATCGAGATCAAAACCTTCGACGAACCTATACCTTATTTGTAT 481

QY 268 TyrValAlaGlyThrValGlyLeuMetSerValProValMetGlyIleAlaProGluSer 287

Db	Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	482	TATGTGCTGTAACGCTTGGTGTATGAGTGTCCAAATTATGAGGATGCCCCGTAATCA						
Qy	288	LysalathrnrhglsuSerValTyrAsnAlaIaleuAlaIeuGlyLeuAlaAsnIleu						
Db	542	AAGCAACAAACAGAGAGCGTATATATATCTGCTTGGCTTCGGGATCGCAATCAATTA						
Qy	308	ThrseuileuAraAspValGlyGluAspAlaAgaArgGlyArgValTyrIleuProGln						
Db	602	ACTACACATACCTCAGAGATGTGGAGAAATGCCAGAAAGAGAAAGTCTACTTGCTCAA						
Qy	328	AspGluLeuAlaGlnAlaGlyLeu 335						
Db	662	GATGATTTATACCCAGCAAGTCTTA 685						
RESULT 9								
BM409200								
LOCUS								
DEFINITION	EST583527 tomato breaker fruit Lycopersicon esculentum cDNA clone	682 bp	mRNA	linear	EST 22-JAN-2002			
ACCESSION	BM409200							
VERSION	BM409200.1	GI:18260830						
KEYWORDS	EST.							
SOURCE	Lycopersicon esculentum (tomato)							
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asterid; Lamid; Solanales; Solanaceae; Solanum; Lycopersicon.							
REFERENCE	1 (bases 1 to 682)							
AUTHORS	Alcala,J., Vrebalov,J., White,R., Vistion,T., Karamycheva,S.A., Tsai,J., Bougri,O., Kirtnes,E., Utterback,T., Van Aken,S., Roming,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.							
TITLE	Generation of ESTs from tomato fruit tissue, breaker stage (2002)							
JOURNAL	Unpublished							
COMMENT	Contact: CUGI Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a> This clone is available through the Clemson University Genomics Institute Seq primer: T3.							
FEATURES								
source	Location/Qualifiers 1..682 /organism="Lycopersicon esculentum" /mol_type="mRNA" /cultivar="TA496" /db_xref="taxon:4081" /clone="CLG4701" /tissue_type="pericarp" /dev_stage="breaker" /lab_host="SOLR" /clone_idb="tomato breaker fruit" /note="Vector: pBlueScriptmCudaapt; Site 1: EcorI; Site 2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopen accumulation on the blossom end of fruit). Fruits were cut in half and the seeds and locules were discarded prior to freezing the pericarp."							
BASE COUNT	217 a 114 c 173 g 178 t							
ORIGIN								
Alignment Scores:								
Prod. No.:	7.81e-115	Length:	682					
Score:	1083.00	Matches:	207					
Percent Similarity:	94.71%	Conservative:	8					
Best Local Similarity:	91.19%	Mismatches:	12					
Query Match:	47.71%	Indels:	0					
Db:	12	Gaps:	0					

Oy	155	AlaGluIuYrAlaIeYrThrPheTyrlEuGIYThrlYsLeuMeCThrProGIuAaYrArg	174
Db	1	GCAGAGrATGCAGAAAGCGTTAACTTAGAACTATGCTTAATGACTCCCGAGAGAAAGG	60
Oy	175	AlaIleTrpAlaIleTyYrValTrpCySaYrAqThzAspGIuLeuValAspGIYProAsn	194
Db	61	GCATATCTGGCGCAATATATGTATGGTGCAGAAAGAACATGAACTGTTGATGGCCAAAC	120
Oy	195	AlaSerHisIleThrProGIuAlaLeuAspArGTTrpGIuThzArGLeuGIuAspIlePhe	214
Db	121	GCATCATATATTACCCCGCAGCGCTTGTAGATAGGGGAGAAATAGGCTTAGAAGATGTTTTC	180
Oy	215	SerGIYArGrProPheAspMetLeuAspAlaIleLeuSerAspThrValSerArgPhePro	234
Db	181	AATGGGGGGCCATTATGCATGCTCGATGGTGGCTTGGCCAGATACAGTTTCACTTCCAA	240
Oy	235	ValAspIleGIeInProPheArGAspMetIleGIuGIYMeArGMeAspLeuTrpLYsSer	254
Db	241	GTTGATATTTCAGCCATTCAAGAGATATGATTAAGAGATGGGTATGGACTTGAGAAATCG	300
Oy	255	ArgTrYrlsThrPheAspGIuLeuTyrlEuTyrlCySerTyrlYrValAlaGIYThrValGIY	274
Db	301	AGATACAAAAAAGCTCGACGAACCTATACCTTATTTGTTATTTATGTTGCTGTACCGATGGG	360
Oy	275	LeuMeSerValProValMeGIYIleAlaProGIuSerLYsAlaThrTrGIuSerVal	294
Db	361	TTGATGTAGGTCTCCAAATTATGGGTATGCCCTCGAATCAAGGCAACACGAGAGCGTA	420
Oy	295	TyrAsnAlaIleLeuAlaIleuAlaLeuAlaAsnGIuLeuThrAsnIleLeuArgAspVal	314
Db	421	TATATATGCTGCTTTGGCTCTGGGATGCGCAATCAATTAACATACTACAGAGATGT	480
Oy	315	GIYGIuAspAlaArGArGGIYArGValTyrlEuProGIuAspGIuLeuAlaGIuAlaGIY	334
Db	481	GGAGAAATGCGCAGAAAGAGAGAGTCTACTTGGCTCAAGATGAATATATCATCAGACGAGT	540
Oy	335	LeuSerAspGIuAspIlePheAlaGIYArGValThzAspLYsTrpAAsnPheMeLYs	354
Db	541	CTATCCAGTAAAGATATATTGCTCGAAGGGTGCACATTAATGAGAAATCTTTATGAAAG	600
Oy	355	LYsGIuIleGIeInArGAlaArGIYsPhePheAspGIuSerGIuLYsGIYValThrGIuLeu	374
Db	601	AAACAAATACCTAGGCGCAAGAAAGTTCTTTGATGAGCAGAGAAAGCGTCACGAATTG	660
Oy	375	AspSerAlaSerArGTTrpPro 381	
Db	661	AGTCACGCTAGTAGATTCCT 681	
RESULT 10			
LOCUS	AM442407	672 bp	mRNA
DEFINITION	EST311803 tomato fruit red ripe, TMU lycopersicon esculentum cDNA		
ACCESSION	AM442407		
VERSION	AM442407.1		
KEYWORDS	GI:6977658		
ORGANISM	EST.		
SOURCE	Lycopersicon esculentum (tomato)		
REFERENCE	1 (bases 1 to 672)		
AUTHORS	Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upson,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Roming,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.		
TITLE	Generation of ESTs from tomato fruit tissue		
JOURNAL	Unpublished		
COMMENT	Contact: CUGI		
	Clemson University Genomics Institute		
	Clemson University		
	100 Jordan Hall, Clemson, SC 29634, USA		
	Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a>		
	5 prime sequence.		

# FEATURES

## source

Location/Qualifiers  
1. 672  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEN221.4"  
/issue\_type="pericarp"  
/dev\_stage="red ripe (7-20 days post-breaker)"  
/clone\_lib="tomato fruit red ripe, TAMU"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI, Site 2: XhoI, supplier: Giovannoni; Fruit were tagged at the breaker stage (first sign of lycopen accumulation on the blossom end of the fruit) and harvested 7 days post-breaker (fully red-ripe), 10 days post breaker, and 20 days post-breaker (over-ripe). 20 day fruit which showed external or internal signs of pathogenesis were discarded. Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

BASE COUNT 212 a 112 c 170 g 178 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1,44e-113 Length: 672  
Score: 1072.00 Matches: 205  
Percent Similarity: 95.07% Conservative: 7  
Best Local Similarity: 91.93% Mismatches: 11  
Query Match: 47.22% Indels: 0  
DB: 9 Gaps: 0

US-09-847-081b-2 (1-440) x AM442407 (1-672)  
QY 160 ThrpetyrLeuGlyThrltysLeuMetThrProGluuArgArgAlaIleTrrPaIaIle 179  
DB 3 ACCTTAACCTTAGAAGCTATGCTATGATCCCGAGAGAAAGAGGCGCTATCGGCAATA 62  
QY 180 TyrValTrrPcyArgArgThraspGluLeuValAspGlyProAsnAlaSerHisIleThr 199  
DB 63 TATGATGATGAGAGAAAGACATGATGCTTGATGGCCAAAGCATCATATATATAC 122  
QY 200 ProGlnAlaLeuAspArgTrrPgluThraGluGluAspIlePheSerGlyArgProPhe 219  
DB 123 CCGGAGACCTTGATGATGAGGAGAAATAGCGTAAAGATGTTTCAATGGCGGCATTT 182  
QY 220 AspMetLeuAspAlaIleLeuSerAspThrValSerArgPheProValAspIleGlnPro 239  
DB 183 GACATGCTGATGATGCTTTCCTTCATGATGCTTTCATGCTTTCATGATGATGATGATG 242  
QY 240 PheArgAspMetIleGluGlyMetArgMetAspLeuTrrPlysSerArgTyrIlyThrPhe 259  
DB 243 TTCAGAGATATGATTAAGGAATGCGTATGACTTGAGAAATCGAGATACAAAACCTTC 302  
QY 260 AspGluLeuTrrLeuTrrCysTrrTyrValAlaGlyThrValGlyLeuMetSerValPro 279  
DB 303 GACGAACATACCTTAATGTTATGTTGCTGATGCGTATGCGTATGATGATGATGATG 362  
QY 280 ValMetGlyIleAlaProGluSerLysAlaThrThrGluSerValTrrAsnAlaIleLeu 299  
DB 363 ATATAGGATGATGCGCTTAATCAAGGCAACAGAGAGCGTATATATATGCTGCTTTG 422  
QY 300 AlaLeuGlyLeuAlaAsnGluLeuThrAsnIleLeuArgAspValGlyLubAspAlaArg 319  
DB 423 GGTCTGGGATCGCAAAATCAATTAATCAATCAATCAATCAATCAATCAATCAATCAAT 482  
QY 320 ArgGlyArgValTrrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAsp 339  
DB 483 AGAGGAAGAGTCTACTTGCCTCAAGATGATTAAGCAGGAGGTCTATCCGATGAAGAT 542  
QY 340 IlePheAlaGlyArgValThrAspLysTrrPargAspPheMetLysLysGlnIleGlnArg 359  
DB 543 ATATTTGCTGAGAAAGGATGACCGATTAATGAGATCTTTATGAGAAACAAATATCATAG 602  
QY 360 AlaArgLysPhePheAspGluSerGlyLysGlyValThrGluLeuAspSerAlaSerArg 379

DB 603 GCAAGAAAGTCTTGTATGAGGAGAGAAAGCGGTACAGATTAAGTCTACCTAGTAGA 662  
QY 380 TrrProVal 382  
DB 663 TTCCCTGTA 671

RESULT 11  
BE433198 642 bp mRNA linear EST 18-MAY-2001  
LOCUS BE433198  
DEFINITION EST399727 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA  
ACCESSION BE433198  
VERSION BE433198.1 GI:9431041  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Bukarjota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 642)  
Alcala,J., Vredalov,J., White,R., van der Heeven,R.S., Holt,I.E., Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley

AUTHORS  
Generation of ESTs from tomato fruit tissue, breaker stage  
S.D.  
Unpublished  
CONTACT: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

TITLE  
JOURNAL  
COMMENT

FEATURES  
source  
Location/Qualifiers  
1. 642  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEB12N5"  
/issue\_type="pericarp"  
/dev\_stage="breaker"  
/lab\_host="SOLR"  
/clone\_lib="tomato breaker fruit, TIGR"  
/note="Vector: pBluescriptSMCtdapt; Site 1: EcoRI, Site 2: XhoI; Fruit were harvested at the breaker stage (first sign of lycopen accumulation on the blossom end of the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

BASE COUNT 194 a 103 c 167 g 178 t  
ORIGIN

Alignment Scores:  
Pred. No.: 6.25e-111 Length: 642  
Score: 1049.00 Matches: 199  
Percent Similarity: 95.77% Conservative: 5  
Best Local Similarity: 93.43% Mismatches: 9  
Query Match: 46.21% Indels: 0  
DB: 10 Gaps: 0

US-09-847-081b-2 (1-440) x BE433198 (1-642)

QY 140 LeuGlyLeuLeuSerGluAlaTrrAspArgCysGlyGluValCysAlaGlyTrrAlaLys 159  
DB 2 TTGGGCTTGTGAGAGGACATATGATGCTGCTGAGATGATGCTGAGATGCTGAGATG 61  
QY 160 ThrpetyrLeuGlyThrltysLeuMetThrProGluuArgArgAlaIleTrrPaIaIle 179  
DB 62 ACCTTAACCTTAGAAGCTATGCTATGATCCCGAGAGAAAGAGGCGCTATCGGCAATA 121  
QY 180 TyrValTrrPcyArgArgThraspGluLeuValAspGlyProAsnAlaSerHisIleThr 199  
DB 122 TATGATGATGAGAGAAAGACATGATGCTTGATGCGCCAAAGCATCATATATATAC 181





SOURCE Solanum tuberosum (potato)  
 ORGANISM Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum.  
 REFERENCE 1 (bases 1 to 677)  
 AUTHORS Crookshanks, M., Emmersen, J., Wellinder, K.G. and Nielsen, K.L.  
 TITLE The potato tuber transcriptome: analysis of 6077 expressed sequence  
 tags  
 JOURNAL FEBS Lett. 506 (2), 123-126 (2001)  
 MEDLINE 21475600  
 PUBMED 11591384  
 COMMENT Contact: Karen G. Wellinder  
 Institut for Biocnologi  
 Aalborg Universitet  
 Sohngaardsklovnvej 49, 9000 Aalborg, Denmark  
 Tel.: +45 96358467  
 Fax: +45 98141808  
 Email: kgy@bio.auc.dk  
 Sequenced from the 5' end.  
 High quality sequence stop: 677  
 PolyA=No.

FEATURES  
 source Location/Qualifiers  
 1..677  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /culturvar="Field grown Kuras"  
 /db\_xref="taxon:4113"  
 /tissue\_type="tuber"  
 /clone\_lib="Mature tuber lambda ZAP"  
 /note="Vector: Lambda ZAP"  
 BASE COUNT 192 a 102 c 198 g 185 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2.5e-105 Length: 677  
 Score: 1001.00 Matches: 195  
 Percent Similarity: 91.15% Conservat: 11  
 Best Local Similarity: 86.28% Mismatches: 20  
 Query Match: 44.10% Indels: 1  
 DB: 10 Gaps: 0

US-09-847-081b-2 (1-440) x BG351357 (1-677)

QY 25 AappserValaArgGluGlyAsnaArgValPheValSerSerArgPheLeuAlaArgAspArg 44  
 Db 1 GATTGAGTCCGAGAGGGAACCGGG-TTGAATCATCCAGGTCCCATCTCCGATAGG 59  
 QY 45 AsnLeuMetTrpAsnGlyAArgIleLysLysGlyGlyAArgGlnArgTrpAsnPhgGlySer 64  
 Db 60 AATTGAGTGTGGAAGGAGGATTCCAAGAAAGGTGGAGACAGAGGTGGAAATTTGGGTTT 119  
 QY 65 LeuIleAlaAspProArgTrpSerCysLeuGlyGlySerArgTrpGluLysGlySerThr 84  
 Db 120 TTAATGCAAGATTGAGATATTCTGTGTAGCAAGATCAAGAACTGAGAAAGAGAGT 179  
 QY 85 PheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThrValSerSerGlu 104  
 Db 180 TTTTGTGACATCCAGTGTGGCTGAGTCCAGCTGGAAGATGGCTGTTCATCAGAG 239  
 QY 105 LysLysValIYAspValValLeuLysGlnAlaAlaLeuValIYAspGlnLeuArgSer 124  
 Db 240 AAAAAGGTGTAGAGGTGTATTGAAGCAGGAGCTTTAGTGAAGAGCATCTGATATT 299  
 QY 125 ThrAspAspLeuGluValIYAspProAspIleValValProGlyAsnLeuGlyLeuLeuSer 144  
 Db 300 ACTGAGCATATGAGAGTGAAGCCGAGATATTGTTCCGGAATTGGGCTTGTAGT 359  
 QY 145 GluAlaIYAspAspArgCysGlyGluValCysAlaGluIYAspAlaIYAspThrPheTrpLeuGly 164  
 Db 360 GAAAGCATATGATCGTGTGGCCAGAGATGTGAGAGTATGCTTAAGCATTTTACTTAGGA 419  
 QY 165 ThrLysLeuMetThrProGluArgArgAlaIleTrpAlaIleTrpValIYAspArg 184

Db 420 ACCATGCTAATGATCCAGACAGAGAAGAGCTATCTGGGCAATATATGTGGTGACAG 479  
 QY 185 ArgTrpAspGluLeuValIYAspGlyProAsnAlaSerHisIleThrProGlnAlaLeuAsp 204  
 Db 480 AGAAGCATATGAGCTGTGTGATGGCCCTTAATGATACACATATCCCAAGCTTTAGT 539  
 QY 205 ArgTrpGluThrArgLeuGluValAspIlePheSerGlyArgProPheAspMetLeuAspAla 224  
 Db 540 AGGTGGAGGCCAGCTCGAAGATATTTTCAACGGGCGCCCATATGATATGCTTATGCA 599  
 QY 225 AlaLeuSerAspTrpThrValSerArgPheProValAlaPileGlnProPheArgAspMetIle 244  
 Db 600 GCTTATCCGATACGTGTTCCAATTTCTGTGCAATATTCAGCCATTCAGATATAGT 659  
 QY 245 GluGlyMetArgMetAsp 250  
 Db 660 GAAAGAAATGCGTATGAC 677

RESULT 14  
 BM412533  
 LOCUS BM412533  
 DEFINITION EST186860 tomato breaker fruit Lycopersicon esculentum cDNA clone  
 ACCESSION BM412533  
 VERSION BM412533.1 GI:18264163  
 KEYWORDS EST.  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
 REFERENCE 1 (bases 1 to 634)  
 AUTHORS Alcalá, J., Vrebalov, J., White, R., Vaison, T., Karanycheva, S.A., Tsai,  
 J., Bougri, O., Kirhness, E., Ulteback, T., Van Aken, S., Roming,  
 C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.  
 TITLE Generation of ESTs from tomato fruit tissue, breaker stage (2002)  
 JOURNAL Unpublished  
 COMMENT Contact: CUGI  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 This clone is available through the Clemson University Genomics  
 Institute  
 Seq primer: T3.  
 FEATURES  
 source Location/Qualifiers  
 1..634  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /culturvar="TA496"  
 /db\_xref="taxon:4081"  
 /clone\_lib="Pericarp"  
 /tissue\_type="Pericarp"  
 /dev\_stage="breaker"  
 /lab\_host="SOLR"  
 /clone\_lib="tomato breaker fruit"  
 /note="Vector: pBluescriptSMcadapt, Site 1, EcoRI,  
 Site 2: XhoI; supplier: Boyce Thompson Institute;  
 sequencing: The Institute for Genomic Research. Fruit  
 were harvested at the breaker stage (first sign of  
 lycopen accumulation on the blossom end of fruit). Fruit  
 were cut in half and the seeds and locules were discarded  
 prior to freezing the pericarp."

BASE COUNT 200 a 105 c 160 g 169 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 6.58e-105 Length: 634  
 Score: 997.00 Matches: 194  
 Percent Similarity: 94.31% Conservat: 5  
 Best Local Similarity: 91.94% Mismatches: 12  
 Query Match: 43.92% Indels: 1

DB: 12 Gaps: 0

US-09-847-081b-2 (1-440) x BM412533 (1-634)

QY 155 AlaGluTyrAlaValThrPheTyrLeuGlyThrLysLeuMetThrProGluArgArg 174

DB 1 GCAGAGATGCAAGAACGTTTAACTAGAACTATGCTAAAGACTCCCGAGAGAAAGG 60

QY 175 AlaIleTrrAlaIleTyrValTrrCysArgArgThrAspGluLeuValAspGlyProAsn 194

DB 61 GCTATTCGGGCAATATATGATGTGCGAGAAACAGATGAACTTTGATGGCCCAAC 120

QY 195 AlaSerHisIleThrProGlnAlaLeuAspArgTrrGluThrArgLeuGluAspIlePhe 214

DB 121 GCATCATATATTTACCCCGGACGCTTATGATGGTGGGAAATAGGCTAGAAAGATGTTTC 180

QY 215 SerGlyArgProPheAspMetLeuAspAlaLeuSerAspThrValSerArgPhePro 234

DB 181 AATGGGGGGCCATTTGACATGCTCGATGCTCTTGTCCGATACAGTTTCTAACTTCCA 240

QY 235 ValAspIleGlnProPheArgAspMetIleGluGluMetArgMetAspLeuTrrPheSer 254

DB 241 GTTATATTCAGCCATTCAGATATGATTAAGAAAGATGCGTATGAGCTTAGAAATCG 300

QY 255 ArgTyrIleThrPheAspGluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGly 274

DB 301 AGATACAAAACCTTCGACGAACTATACCTTATTTGTTATTTGTTGCTGGTACGGTTGG 360

QY 275 LeuMetSerValProValMetGlyIleAlaProGluSerLysAlaThrThrGluSerVal 294

DB 361 TTGATGAGTGTTCATATATGATGCTATCCCGCTGATCAAGCAACAGAGAGCGCTA 420

QY 295 TyrAsnAlaAlaLeuAlaLeuGlyLeuAlaAsnGluLeuThrAsnIleLeuArgAspVal 314

DB 421 TATATGCTGCTTGGCTCTGGGATGCGAAATCAATTAACATTAATCAATCAAGATGT 480

QY 315 GlyLysAlaAlaArgArgGlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGly 334

DB 481 GGAAGAAGATGCGAAGAGAGAGAGTCTACTTGCCTCAAGATGAAATTGACAGCGAGGT 540

QY 335 LeuSerAspGluAspIlePheAlaGlyArgValThrAspLysTrrPheAsnPheMetLys 354

DB 541 CTATCCGATGAAGATATTTTGTGCGAAGCGGTGACCGATAATGAGAAATCTTTATGAG 600

QY 355 LysGlnIleGlnArgAlaArgLysPhePheAsp 365

DB 601 AAACAATA-CATYAGGGGCGAGAAAGTTCTTGAT 632

RESULT 15

CA999095 704 bp mRNA linear EST 10-JAN-2003

LOCUS S234N\_H03 Rice cold stress germination cDNA library Oryza sativa

DEFINITION (japonica cultivar-group) cDNA 5', mRNA sequence.

ACCESSION CA999095

VERSION CA999095.1 GI:27576401

KEYWORDS EST.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)

REFERENCE de los Reyes, B.G., Moray, M., Gibbons, J., Varma, T.S.N., Antoine, W., Redus, M., McGrath, J.M., and Halgren, R., Development of a chilling stress EST library of germinating rice (Oryza sativa L. subsp. japonica) enriched with stress-related and novel genes

JOURNAL Unpublished

CONTACT: Benildo G. de los Reyes

Plant Genomics Lab., Department of Crop, Soil and Environmental Sciences

University of Arkansas

115 Plant Science Building, Fayetteville, AR 72701, USA

FEATURES

source

1. 704

Location/Qualifiers

1. 704

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="P1560247"

/db\_xref="taxon:39947"

/issue\_type="coleoptile, radicle, prophyll, immature leaf"

/dev\_stage="embryo at different stages of germination and shoots of germinated seeds under chilling stress (130C/100C)"

/lab\_host="SoliR"

/clone\_lib="rice cold stress germination cDNA library"

/note="Organ: seedlings; Vector: Lambda Uni-Zap XR excised as pBluescript; Site 1: EcoRI; Site 2: XhoI; The cDNAs were derived from reverse transcription of mRNA samples from seeds at different stages of germination and seedlings at early phase of growth under chilling stress (130C/100C). The mRNA pool was used as template for double stranded cDNA synthesis using the Stratagene Uni-Zap XR cDNA synthesis and library kit. A total of 150,000 phages were excised from the primary library as pBluescript phagemid clones. Enrichment of the primary excised library with chilling-induced transcripts was performed by hybridizing the primary excised library colony lifts with the PCR-select subtraction product, with cold germinated cDNA as tester and control temperature-germinated cDNA as driver."

BASE COUNT 197 a 150 c 194 g 162 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 2,48e-103 Length: 704

Score: 984.00 Matches: 183

Percent Similarity: 92.48% Conservative: 26

Best Local Similarity: 80.97% Mismatches: 17

Query Match: 43.35% Indels: 0

Gaps: 14

US-09-847-081b-2 (1-440) x CA999095 (1-704)

QY 203 LeuAspArgTrrGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeu 222

DB 3 CTTCATCGATGGGAGAGATTAAGAGATCTCTTCAGAGCGAGCCCATATGATATGAT 62

QY 223 AspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArgAsp 242

DB 63 GATCGACCCCTCTCGGACACAGTGTCAAGTTTCCAGTATTCACGCCATTCAGAGAC 122

QY 243 MetIleGluGlyMetArgMetAspLeuTrrPheSerArgTrrGluThrPheAspGluLeu 262

DB 123 ATGATTAGAGGAATGAGGCTTGACCTGTGGAATCAAGTATAGAGCTTTGATGAGCTC 182

QY 263 TyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGly 282

DB 183 TACCTTACCTGCTACTACGTTGCTGCGACCGTGTGCTTCATGACAGTACCGTATGGGG 242

QY 283 IleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGly 302

DB 243 ATTGCCCCCGATCGAAGGCTCAACCGAGCGTATCAACGCTGCGCTAGCTTGGG 302

QY 303 LeuAlaAsnGluLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGlyArg 322

DB 303 ATGCCCAACAGCTGAGAGATTTCTCAGAGACGTAGCGCAAGCTCAAGAGAGGGAAGA 362

QY 323 ValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAla 342

DB 363 ATCTACCTTCATGATGATGATGAGAGAGCGAGCGAGCTGACAGAAAGAACTATTTGAGA 422

```
QY 343 GlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArgAlaArgLys 362
Db 423 GGGAAAGTGAATGATTAATGAGAGAGTTCATGAGGACAAATCTCGCGTCCAGGTTA 482
QY 363 PhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpProVal 382
Db 483 TTCTTTGATGAGCGGAGAGGGCGTTCGCGCATCTAGACTCTGCGAGTAGATGGCCGGTT 542
QY 383 LeuThrAlaLeuLeuLeuLysTrpArgLysIleLeuAspGluIleGluAlaAsnAspTyrAsn 402
Db 543 CTGGCATCTTTGTGTTATACCGGACAGATCTTGATGCTATCGAAGCAACGACTACMAC 602
QY 403 AsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIleAla 422
Db 603 AACTTCACCAAGCGCGGTATGTAAACAAGCAAGAGCTGCTGTCTTTACCGGTGCGCT 662
QY 423 TyrAlaLysSerLeuVal 428
Db 663 TATGCAAGAGCGGCGAGTT 680
```

Search completed: January 16, 2004, 08:47:18  
Job time : 2829 secs

